“Think of it as Money”: A History of the VISA Payment System, 1970–1984

David L Stearns

Submitted for the Degree of PhD by Research
University of Edinburgh
August 2007

Abstract

This dissertation is a historical case study of the payment system designed, built, and operated by Visa International Services Association (VISA, hereafter "Visa"). The system is analyzed as a sociotechnical one, consisting of both social and technical elements that mutually constitute and shape one another. The historical narrative concentrates on the period of 1970 to 1984, which roughly corresponds to the tenure of the system's founder and first CEO, Dee Ward Hock. It also focuses primarily upon the events that took place within the United States.

After establishing a theoretical and historical context, I describe why and how the organization now known as Visa was formed. I then explain how the founder and his staff transformed the disintegrated, paper-based credit card systems of the 1960s into the unified, electronic value exchange system we know today. Special attention is paid throughout this narrative to the ways in which the technologies were shaped by political, legal, economic, and cultural forces, as well as the ways in which the system began to alter those social relations in return. In the final chapter, I offer three small extensions to the literature on payment systems, cooperative networks, and technology and culture.

Keywords: VISA, BankAmericard, National BankAmericard Incorporated, Dee Hock, credit cards, payment systems, sociotechnical systems.

Declarations

I, David L Stearns, hereby certify that (a) this thesis has been composed by myself; (b) the work is my own; and (c) the work has not been submitted for any other degree or professional qualification.

Date:

Signature:

Contents

For Chelle

Some Notes on Style

This dissertation was typeset in 11 point Roman using the LaTeX system on a Macintosh PowerBook G4. Citations and the Bibliography were produced automatically using the "jurabib" package according to its "oxford" style. In general, this is similar to the Turabian style, except that the formats used in the citations are identical to those used in the bibliography.

The first citation of a source will be in a fully-specified form, and subsequent citations will be in a shortened form. Repeated citations on the same page will use "Ibid." where appropriate. A full bibliography is included at the end of the dissertation.

Acronyms will be spelled out on first use, except those that are so commonplace as to not require definition (e.g., ATM). A list of acronyms and their meanings is provided at the end.

I will use American spellings throughout this dissertation, with one notable exception: when referring to the paper payment instrument known as a "check" in America, I will use the more international spelling of "cheque" to avoid confusion. Similarly, I will use the international spelling for "travellers cheque."

Acknowledgments

This dissertation would not have been possible without the help of a great many people. First and foremost among them is my supervisor, Donald MacKenzie, who shepherded me through this entire process, providing detailed and helpful feedback all along the way. I would also like to thank Luciana D'Adderio, who served as my secondary supervisor and helped me come to grips with organizational theory. Many thanks as well to the members of my proposal review board: Robin Williams, Stuart Anderson, and John Henry.

I especially want to thank all those who graciously gave of their time to educate me on their experiences with the Visa system. A full list of interviewees is included at the end of the dissertation, but the following people gave even more of their time to read and make comments on the chapters: Tom Cleveland, Walt Conway, Win Derman, Frank Fojtik, Scott Harrison, Tom Honey, Roger Peirce, Bill Powar, Chuck Russell, Ron Schmidt, and B Ray Traweek. Special thanks must also be extended to Ingrid Kollmann, the coordinator of Visa's unofficial alumni network, who connected me with most everybody.

I would also like to thank Cathryn Carson, director of the Office for the History of Science and Technology at University of California, Berkeley, for hosting me as a visiting student researcher. My connection with UC Berkeley opened many doors, and their libraries provided most everything I needed.

My time here in Berkeley was greatly enriched by my conversations with Elihu Gerson, who kindly reviewed my chapters, providing valuable feedback. Without his consistent encouragement, I may not have made it through the most challenging of times.

Many thanks to my family members and friends who took the time to read my chapters, especially my grandmother, Helen Wood. A very special thanks goes to Sue Peterson, who could have been a copy editor and finder of arcane financial information if she was not such a maven at financial advising.

Last, but certainly not least, I wish to thank my wonderful wife Chelle, without whom I would have never attempted this, much less finished it.

Part I

Introduction and Context

Chapter 1

Introduction

> Today, before any audience in the world, I can hold a Visa card overhead and ask, "How many of you recognize this?" Every hand in the room will go up. When I ask, "How many of you can tell me who owns it, how it's governed, or where to buy shares?" a dead silence comes over the room. Something incredible happened, but what, and how?[1]

This quote, which opens the autobiography of Visa's founder, Dee Ward Hock, nicely captures the mystery that is Visa. Although Visa is one of the most well-recognized brands in the world, few people have any idea what Visa is, how it is organized, or how it functions. Even fewer know how it came to be.

This is somewhat surprising considering that Visa and other payment card systems are becoming an increasingly integral part of our daily lives. Millions of merchants around the world now accept these small bits of plastic in lieu of cash or cheques, primarily because those merchants can obtain an authorization, which is a *guarantee of payment*, within a few seconds, regardless of where the card came from, what time it is there, and what currency unit is used for the account.[2] Forty years ago, this was scarcely a dream; today, we hardly think twice about it. How does my Visa card work? And how has it come to feel so "normal," or even "natural?"

This dissertation recaptures that story. It explains the origins of the organization and payment system now known as Visa. It describes how Visa's founder and his staff transformed the disintegrated, paper-based credit card systems of the 1960s into the unified, electronic value exchange system we know today. It examines the ways in which this system was shaped by political, legal, economic, and cultural forces, as well as the ways in which the system began to alter those social relations in return.

[1] Dee Hock, *One From Many: VISA and the Rise of Chaordic Organization* (San Francisco: Berrett-Koehler, 2005), p. xv.

[2] Although I will use American spellings throughout this dissertation, I will use the international spelling for "cheque" instead of the American form of "check" in order to avoid confusion. Similarly, I will use the international spelling for "travellers cheque."

1.1 What is Visa?

In general, Visa is not what you might think it is. Visa is not a part of the US Federal Reserve System or any other central bank. It is not an agency of any government, nor is it a public utility. Visa itself is not even a bank. It does not issue cards, nor does it recruit new merchants to accept those cards. It does not grant credit to consumers, maintain their accounts, nor control the interest rates charged on unpaid balances. It does not even manufacture, distribute, or install those electronic point of sale terminals. So what is it, and what does it actually do?

Visa is a cooperative membership association that is jointly owned by more than 20,000 financial organizations around the world. Those member organizations, which are mostly banks, are both the owners and the customers of Visa. The members compete with one another to issue cards and recruit merchants, but Visa provides them a framework in which they can cooperate to provide a ubiquitous payment service that none could have realistically supplied alone.

The central Visa organization, which consists of only 6,000 employees, serves three core functions. First, they maintain, interpret, and enforce the rules by which transactions are processed and disputes are resolved. Second, they build and operate the worldwide computer systems and telecommunication networks that authorize, clear and settle those transactions in electronic form. Third, they promote the Visa name and service marks through general advertising and commercial sponsorships.

Rising from a minority position in the early 1970s, Visa has since become the largest and most dominant payment card system in the world. There are currently over 1.5 billion plastic cards containing the Visa mark in circulation, and they are honored at millions of merchants and ATMs in 170 different countries. Visa's computer systems authorize thousands of transactions *each second*, and clear and settle upwards of 120 million transactions each night. In 2006, Visa cards were used to purchase $4.6 *trillion* (USD) worth of goods and services worldwide, an increase of 16 percent from the previous year.[3]

1.2 Scope of the Project

This project is primarily a historical case study of a sociotechnical system, that is, one composed of both "social" and "technical" elements.[4] The computer systems, telecommunication networks, and other electronic devices that process Visa transactions are certainly critical to the operation and success of the payment system, but they are only part of what makes Visa work. Without the organizational structure and the operating regulations, the payment system would lack the mechanisms necessary for the competitive members to cooperate. To understand the

[3] Visa, 'VisaNet Fact Sheet' ⟨URL: `http://corporate.visa.com/md/fs/corporate/visanet.jsp`⟩ – visited on 26 June 2007; Visa, 'Visa International Annual Report' (2006), ⟨URL: `http://corporate.visa.com/av/reports/corp_report.jsp`⟩.

[4] Thomas P Hughes, *Networks of Power: Electrification in Western Society, 1880-1930* (Baltimore, MD: Johns Hopkins University Press, 1983).

Visa payment system as a whole, we must embrace both the technical and social dynamics and examine how they influence and shape one another.

My focus in the following chapters will be on recapturing as accurate a history of Visa as possible, so that we have a firm, empirical basis from which we can interact with the existing theories found in the Science and Technology Studies (STS) literature. Throughout the chapters I will draw upon those theories to explain and analyze specific events, and in the conclusion, I will offer three general lessons garnered from the case.

My original research focuses on the period of 1970 through 1984, which corresponds to the formation of the organization through the departure of its founder. This was a particularly prolific period of innovation for Visa. In 1970, the system was entirely paper-based and processed little volume outside the US. By 1984, the system was almost entirely electronic, and truly worldwide in scope. Although many important developments in the Visa system have occurred since 1984, they are not covered by this dissertation, primarily due to time and length limitations.

Additionally, this dissertation focuses primarily upon the events that took place within the United States (US). Although there were licensees outside the US beginning in 1966, the International membership organization was not officially formed until 1974, and even then, the central management and technological development was still based entirely in the US. The specific ways in which the system extended into other countries would provide interesting comparative insights, but they are unfortunately beyond the scope of this dissertation.

1.3 Structure of the Dissertation

This dissertation is divided into two main parts, the first providing a context for understanding the detailed history presented in the second. In general, the first part relies upon existing secondary literature, while the second part is based almost entirely on primary research.

In Chapter 2, I begin to construct the context by reviewing the existing literature that is relevant to the case, establishing that no history has yet been written of Visa from an STS perspective. I then clarify what that perspective is, and how it would generate new insights into the case. I end with a discussion of the research methods I used for the project, their known problems and limitations, and the ways in which I sought to overcome those problems as much as possible.

In Chapter 3, I complete the context by introducing the terms and concepts used in payment systems theory, and reviewing the important innovations that occurred within the American payments industry directly leading up to the formation of the organization now known as Visa. Central to these innovations was the BankAmericard, a general-purpose credit card introduced by Bank of America in 1958, and licensed to other banks beginning in 1966.

In Chapter 4, I begin the detailed history by discussing the problems facing the BankAmericard licensing system in the late 1960s, arguing that a new kind of organization was necessary

to correct them. I then introduce Dee Hock, the designer and first president/CEO of that new organization, which was initially called National BankAmericard Incorporated (NBI). I present Hock's insights into the nature of money and his theories on organizations, both of which influenced his design for NBI, as well as his plans for a fully-electronic value exchange system. I end by recounting the organization's first steps, which were designed to refocus the members on the system as a whole.

In Chapters 5 and 6, I describe how NBI solved the most pressing of the operational problems outlined in Chapter 4: authorization; and clearing and settlement. In Chapter 5, I recount how NBI and others automated the authorization process, and in Chapter 6, I explain how NBI developed their electronic clearing and settlement system. Chapter 6 also chronicles NBI's first significant failure: a complex software system designed for the member banks that proved to be impractical.

In Chapter 7, I examine the various ways in which the payment system was expanded throughout the rest of the 1970s, both in terms of the organization and the computer systems. Although the organizational expansions, which included changing the name to Visa, certainly triggered much of their explosive growth during that period, the expansions to the computer systems enabled them to handle the corresponding increase in transaction load without suffering major system outages.

In Chapter 8, I review the last area of the payment system to become fully automated: the point of sale (POS). Visa accomplished this by not only stimulating the development and adoption of inexpensive POS dial terminals, but also establishing a common standard for encoding the card. Once the terminals were widely adopted, the Visa system processed most transactions entirely in electronic form, approaching Hock's initial vision of a fully-electronic value exchange system.

In Chapter 9, I complete the detailed historical narrative by recounting how the roles played by Hock and the central Visa organization had to be negotiated and worked out through a series of power struggles. Hock had a much larger concept of his role than the members did, and this led to a number of conflicts about what Visa should and should not be doing. These conflicts eventually weakened Hock's influence over the system, ultimately leading to his departure in 1984.

In Chapter 10, I reflect upon the case, deriving three lessons that extend the existing literature on payment systems, cooperative networks of competitors, and networks and social boundaries. The first lesson concerns the seemingly insignificant technology of a "mark," and how it interacts with authority and value-exchange rules in the development and adoption of most payment systems. The second investigates how operating rules might successfully be formed and maintained within a cooperative network of competitors, and how these networks must develop a capacity for resolving disputes about their cooperative work, as well as the interpretation of those rules. The third delves into the ways in which transactional networks interact with existing social boundaries, and how the design of the network and gateways can

either reinforce or undermine those boundaries.

The story of Visa provides us a fascinating example of how a large sociotechnical system originated, automated, and evolved into a something hundreds of millions of people rely upon every day. However, before we can fully appreciate and understand this story, we must first establish a theoretical and historical context. It is to the former part of that context that we now turn.

Chapter 2

Theoretical Context: Literature Review and Methodology

In the previous chapter, I introduced this project and provided a brief orientation to the Visa organization. The purpose of Part I is to provide a context for understanding the detailed history presented in Part II, and this chapter will supply the theoretical side of that. I will begin by reviewing the existing literature that is relevant to this case, arguing that no history of Visa has yet been written from a Science and Technology Studies perspective. I will then clarify what that perspective is, and how it would generate new insights into the case. I will end with a discussion of the research methods I used for this project, their known problems and limitations, and the ways in which I sought to overcome those problems as much as possible.

2.1 Literature Review

Although payment card systems like Visa have become an indispensable part of our everyday lives, they have been investigated by a surprisingly small number of authors. These authors have written from a variety of different perspectives, and one useful way to review the existing secondary sources is by the standpoint the authors take.

2.1.1 Visa and the Payment Card Industry

Writing from the perspective of a commissioned corporate biographer, Chutkow offers the only source solely dedicated to Visa.[1] As with many corporate biographies, this book is formatted for the coffee table, with full-page, glossy images interspersed with text. This is Visa's official corporate history, and as such, can be seen both as a secondary and a primary source; although Chutkow's book discusses the organization's formation and growth, it also reveals much about

[1] Paul Chutkow, *Visa: The Power of an Idea* (Chicago: Harcourt, 2001).

how the current organization wants to be viewed by the public and its employees.[2] His account is broad in scope, and is perhaps the most complete source on Visa's brand development and advertising efforts. Although one of the sections is dedicated to the technological aspects of the system, it is unfortunately thin and lacking in depth.

Two authors have written about payment cards and the banking industry from a journalistic perspective. The most valuable is Nocera's account, which presents Visa as merely one episode in the larger revolution that took place in the American post-war financial industry.[3] Nocera explains how innovations such as the credit card, the mutual fund, and the cash management account gave the middle class access to financial tools and markets that had previously been available only to the money class. Another interesting account is provided by Mayer, who tells numerous stories about how the banking industry has changed since the introduction of electronics, but rarely connects these together into a cohesive argument.[4] Although these sources are informed by the experience the authors gained covering the financial industry for many years, they tend to focus solely on the business aspects of the system, presenting only surface-level descriptions of the technologies that make the system possible.

Writing from the perspective of cultural criticism, Manning discusses credit cards as one method the banks have used to "hook" American consumers on credit, much as a drug pusher uses free samples to entice new users.[5] Manning is concerned that Americans have become too dependent on credit for their basic survival, and blames the credit card issuers for encouraging consumers to live beyond their means. Galanoy, who served as Visa's Director of Communications in the 1970s, shares Manning's concern, but takes a more sinister view.[6] He portrays credit cards as part of a greater conspiracy by the banks to control our lives through an all-encompassing information system he calls "Lifebank." While these authors convey important warnings against the abuse of credit, they tend to conflate the card issuers with the central Visa organization, assuming that they share the same interests. As we shall see, this was often not the case.

On the academic side, a few authors have discussed payment cards in the broader context of payment systems theory. Humphrey provides a clearly-written introduction to the topic of payment systems, and discusses payment cards as one type of non-cash payment option currently favored in developed nations.[7] Fry et al. discuss the importance of reliable payment

[2]The book had only one printing, so it is mostly owned by employees. For the role this kind of "saga" plays in forming organizational identity, see Burton R Clark, 'The Organizational Saga in Higher Education', *Administrative Science Quarterly* 17 (1972).

[3]Joseph Nocera, *A Piece of the Action: How the Middle Class Joined the Money Class* (New York: Simon and Schuster, 1994).

[4]Martin Mayer, *The Bankers: The Next Generation* (New York: Truman Talley, 1997).

[5]Robert D Manning, *Credit Card Nation: The Consequences of America's Addiction to Credit* (New York: Basic Books, 2000).

[6]Terry Galanoy, *Charge It: Inside the Credit Card Conspiracy* (New York: GP Putnam's Sons, 1980).

[7]David B Humphrey, *Payment Systems: Principles, Practice and Improvements* (Washington DC: The World Bank, 1995).

systems in market economies, and chart some of the current structural and behavioral trends.[8] They note that cash is still the predominant payment system, especially in poorer countries, but richer nations are quickly shifting their low-value payments to card-based electronic systems. Hopton discusses the various approaches to pricing a payment system, noting that the marginal cost pricing used for utilities is not appropriate for payment systems, as they must compete with all other existing forms of payment, each of which has different profitability constraints.[9] These authors provide a much needed analysis of the issues and questions surrounding any payment system, but do not offer any detailed history of Visa itself, nor any analysis of how the technologies involved shaped the payment system.

A number of authors have written about Visa and payment cards from the perspectives of business, banking, and finance. Dougherty's Harvard Business School case study of Visa in 1981 provides an exceedingly useful snapshot of the organization as it existed at that time.[10] Mandell, a financial historian, offers what seems to be the first comprehensive history of the American payment card industry.[11] His focus is almost entirely upon the experiences of the issuing and acquiring banks, though he does include two chapters on some of the technological issues faced by electronic payment card systems. Kirkman wrote from a similar perspective about the development of electronic funds transfer (EFT) systems in the UK, as did Frazer.[12] Wonglimpiyarat examines payment cards from the related perspective of technology management, detailing the business strategies taken by various firms in the UK payments industry.[13] Although some of these authors discuss Visa in particular, their analyses focus almost entirely on the business strategy dynamics, and do not take into account the way the technologies shape the system, nor the ways in which the business strategies shape the technologies.

Howels and Hine are the noted exception to this.[14] They recount the failed attempt to create a single, shared, national EFT system in the UK, and theirs is the only book that delves into the technological issues, showing how they played a major role in the effort's demise. Although it is not about the Visa system in particular, it does provide an interesting counterexample. The EFTPOS UK organization was similar to Visa, but instead of building a system that was good-enough for the present needs, they designed a "best-of-breed" system that turned out to be too complicated and expensive to build.

Lastly, a few authors have written about the payment card industry from an economic per-

[8] Maxwell J Fry et al., *Payment Systems in Global Perspective* (London: Routledge, 1999).

[9] David Hopton, *Payment Systems: A Case for Consensus* (Bank for International Settlements, 1983).

[10] J Stewart Dougherty, 'Visa International: The Management of Change', Harvard Business School Case Study (1981).

[11] Lewis Mandell, *The Credit Card Industry: A History* (Boston: Twayne, 1990).

[12] Patrick Kirkman, *Electronic Funds Transfer Systems: The Revolution in Cashless Banking and Payment Methods* (Oxford: Basil Blackwell, 1987); Patrick Frazer, *Plastic and Electronic Money: New Payment Systems and Their Implications* (Cambridge: Woodhead-Faulkner, 1985).

[13] Jarunee Wonglimpiyarat, *Strategies of Competition in the Bank Card Business: Innovation Management in a Complex Economic Environment* (Brighton: Sussex Academic Press, 2004).

[14] John Howells and James Hine, editors, *Innovative Banking: Competition and the Management of a New Networks Technology* (London: Routledge, 1993).

spective. Evans and Schmalensee are perhaps the most well-known, and their original 1999 book has now been superseded by a second edition based upon the theories of multi-sided platform economics.[15] Evans and Schmalensee use these theories to explain the classic "chicken-and-egg" problem facing any new payment system: consumers will use it only if enough merchants accept it, and merchants will accept it only if enough consumers want to use it. To create a sustainable system, providers need to price their service so that it will be attractive to both sides, and this rarely works out to be an even distribution. Just as singles bars and dating services charge men more than women, the payment card networks typically charge merchants more than consumers on a per-transaction basis.[16] Other economic analyses of the credit card and ATM networks can also be found in Russell, McLeod, and Hayashi et al.[17]

While these various perspectives are helpful for understanding the complex dynamics that shaped the history of Visa and the payment card industry, one important perspective is missing: Science and Technology Studies (STS). The STS perspective attempts to bring together a deep understanding of the technical dynamics operating in any historical case with a careful examination of the economic, political, organizational, legal, and cultural dynamics. Only when we consider all these in concert can we fully understand why a given technological system originated and evolved the way it did. The STS perspective leads us to ask different kinds of questions, the answers to which provide us with new insights into the case.

2.1.2 Science and Technology Studies Perspectives

Given that a history of Visa written from an STS perspective would likely generate new insights, I should now explain what I mean by an "STS perspective." STS is a wide and diverse field, and its participants have developed many different frameworks and approaches for studying the history of technology. However, two related approaches have proven to be particularly useful for analyzing historical cases such as Visa: sociotechnical systems, and actor-network theory. Although these two approaches are similar, they have subtle differences that will be explained in the following sections.

Sociotechnical Systems

The sociotechnical systems approach, most notably developed by Hughes, views material artifacts as components of larger systems of production and use that involve both technical and

[15] David Evans and Richard Schmalensee, *Paying with Plastic: The Digital Revolution in Buying and Borrowing*, 2nd edition (Cambridge, MA: MIT Press, 2005).

[16] Of course, consumers that incur finance changes pay much more over the long-run.

[17] Thomas Russell, *The Economics of Bank Credit Cards* (New York: Praeger, 1975); Robert W McLeod, *Bank Credit Cards for EFTS: A Cost-Benefit Analysis* (Ann Arbor, MI: University Microfilms International, 1979); Fumiko Hayashi, Richard Sullivan and Stuart E Weiner, *A Guide to the ATM and Debit Card Industry* (Kansas City, MO: Payment Systems Research Department, Federal Reserve Bank of Kansas City, 2003).

social elements.[18] The language here comes from General System Theory, where a *system* is defined as a dynamic combination of various *elements in relation*, each of which can be another, nested system.[19] A system is also more than just the sum of these parts: the interacting elements of a system often produce new *emergent* properties that cannot be said to originate solely from any one element. Thus, a system is essentially irreducible, as its behavior cannot be explained by examining its elements in isolation. In order to understand a system fully, one must consider the interactions *between* elements in addition to the elements themselves.

A *sociotechnical system* then is comprised of both technical and social elements interacting together to produce something more than the sum of its parts. The technical elements are the typical objects of inquiry for the traditional historian of technology: material artifacts, techniques, devices, machines, etc. The social elements consist of humans and the various products of their social relations: organizations, economic markets, legal codes, governmental regulations, cultural values, and political power. To understand the technical elements, the modern history of technology suggests that one must also consider the social ones, and the ways in which they relate to one another. Through the dynamic interaction of all these elements, the system is constantly brought into being and refashioned, producing the historical events that historians attempt to explain.[20]

The relationship and ultimate distinction between the "social" and "technical" elements within a system has been debated and softened over the years. Initially, authors tended to treat them as two separate spheres that merely interacted, and this led the authors to emphasize too strongly the ways in which material artifacts are shaped by social relations, neglecting the ways those artifacts influence, or even enable, those social relations in return.[21] The first (1985) edition of MacKenzie and Wajcman's edited collection on the Social Shaping of Technology (SST) made this mistake, as did the early work by Pinch and Bijker on their Social Construction of Technological Systems (SCOT) framework.[22] The SCOT framework posits that relevant social

[18]This was originally formulated in Hughes, *Networks of Power*. See also Thomas P Hughes, 'The Evolution of Large Technological Systems', in: Wiebe E Bijker, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987); and Thomas P Hughes, 'The Seamless Web: Technology, Science, Etcetera, Etcetera', *Social Studies of Science* 16:2 (May 1986). For a review of the use of the systems perspective in *Technology and Culture* articles, see John M Staudenmaier, *Technology's Storytellers: Reweaving the Human Fabric* (Cambridge, MA: MIT Press, 1985). Staudenmaier notes that the systems perspective can devolve into a highly functional approach, ignoring individual creativity, personality and sociological factors. However, he praises Hughes's blending of the systems perspective with technological traditions and sociological analysis, which is what I am here referring to as the "sociotechnical systems" approach.

[19]Ludwig von Bertalanffy, *General System Theory: Foundations, Development, Applications* (New York: George Braziller, 1968); Gerald M Weinberg, *An Introduction to General Systems Thinking* (New York: John Wiley and Sons, 1975).

[20]Hughes, *Networks of Power*.

[21]Donald MacKenzie and Judy Wajcman, editors, *The Social Shaping of Technology*, 2nd edition (1985; reprint, Maidenhead: Open University Press, 1999), pp. 22–23.

[22]Trevor J Pinch and Wiebe E Bijker, 'The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other', *Social Studies of Science* 14:3 (August 1984). An updated and shortened version is in Trevor J Pinch and Wiebe E Bijker, 'The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other', in:

groups shape variants of artifacts towards solutions to their constructed problems. Although SCOT admirably demonstrates how groups can shape the design of the artifacts, there is little consideration for how the artifacts alter the way those groups form, interact, and construct their problems in return.

MacKenzie began to challenge the distinction between the "technical" and the "political" in his study of nuclear missile guidance systems, concluding that "it is too weak a position to even see technology and politics as interacting: there is no categorical distinction to be made between the two."[23] From an analytical standpoint, it is difficult to categorize elements strictly as "technical" or "social" because the very knowledge we create about material artifacts is always shaped, at least to some degree, by so-called social forces. Participants may construct separate categories for themselves for specific reasons (such as ensuring a division of labor), but these should never be taken as "natural" or given by the analyst. Latour (an actor-network theorist) further challenged the idea that "society" somehow exists outside or separate from our technologies, arguing instead that both society and technologies are continuously re-created through "institutionalized transactions."[24] In the second edition of their edited collection on SST, MacKenzie and Wajcman reformulated their theoretical stance, stressing that "technology and society are mutual constitutive."[25]

If the distinction between elements is often fuzzy, establishing the boundaries of the sociotechnical system can also be difficult. Because systems are hierarchical, any given system is also in constant interaction as an element within another, encompassing system. This leads to the methodological problem of deciding where the line is between the system's content and its context. One solution, suggested by Hughes, is to make a distinction between interaction and control.[26] By focussing on the *system builders* (those striving to bring the system into existence and shape its evolution), the historian can determine the system's boundary by what the builders are able to control; anything within that control is part of the system's content, while anything outside is part of the context. The context and content still interact and influence one another, but the system builders have no direct control over the elements in the context.

For Hughes, focussing on the system builders not only helps to define the system's boundary, but also enables the historian to speak about a specific, temporary goal towards which the system is moving. This is perhaps an important weakness in his approach, as it tends to ne-

Wiebe E Bijker, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987).

[23]Donald MacKenzie, *Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance* (Cambridge, MA: MIT Press, 1990), pp. 412–413.

[24]Bruno Latour, 'Ethnography of a High-Tech Case: About Aramis', in: Pierre Lemonnier, editor, *Technological Choices: Transformation in Material Cultures Since the Neolithic* (London: Routledge, 1993), p. 395. See also Bruno Latour, *Aramis, or, The Love of Technology* (Cambridge, MA: Harvard University Press, 1996); Bruno Latour, *Reassembling the Social : An Introduction to Actor-Network-Theory* (Oxford: Oxford University Press, 2005); and Shirley Strum and Bruno Latour, 'Redefining the Social Link: From Baboon to Humans', *Social Science Information* 26:4 (1987).

[25]MacKenzie and Wajcman, p. 23.

[26]Hughes, 'Seamless Web'.

glect the other groups in the system that may have different ideas about what the goal should be. Nevertheless, a key part of Hughes's framework is the idea of the *reverse salient*, which is something hampering the system's progress towards the builders' current goal.[27] The system builders remove a reverse salient by defining one or more critical problems, the articulation of which often implies particular solutions. Hughes found these concepts useful in describing Edison's work on the filament, but MacKenzie warns that determining goals, reverse salients, and critical problems, is a process of knowledge creation, and thus should never be assumed as "given, independent of the actors involved." The system builders may not all agree on what goal should be obtained, what is hampering progress toward it, and how it might be solved. Causation may also run the other way—actors may formulate critical problems based on what they can actually accomplish, and then argue for a corresponding reverse salient.[28]

Another important aspect of the sociotechnical systems approach is viewing the system builders as *heterogeneous engineers*, that is, people who solve economic, legal, political, or organizational problems in addition to technical ones.[29] System builders are often generalist entrepreneurs who are determined to bring about change, and inventing new technologies is only one of many ways in which they accomplish that. They rarely segregate problems into just one category, and even the creation of a single artifact, such as a light bulb filament, might be seen as solving both an economic as well as a technical problem.[30]

Actor-Network Theory

The second STS approach to review is actor-network theory (ANT), which was initially defined by Callon and Latour.[31] ANT begins with an attempt to eliminate the problematic distinctions between "technical" and "social" elements discussed earlier, and as such, refers to both inanimate artifacts and animate humans as *actors*. Actors are assumed to have interests, which they follow until a network builder attempts to *enroll* them into a new network. This enrollment is accomplished through a *trial of strength*, where the network builder *translates*, or literally

[27] Hughes, *Networks of Power*.

[28] Donald MacKenzie, 'Missile Accuracy: A Case Study in the Social Process of Technological Change', in: Weibe E Bijker, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987), pp. 197–199.

[29] Although this term is most commonly attributed to the actor-network theorist, John Law, Hughes uses similar language when describing the actions of system builders. See John Law, 'Technology and Heterogeneous Engineering: The Case of Portuguese Expansion', in: Wiebe E Bijker, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987); and Hughes, 'Seamless Web'.

[30] Hughes, *Networks of Power*.

[31] Michel Callon, 'The Sociology of an Actor-Network: The Case of the Electric Vehicle', in: Michel Callon, John Law and Arie Rip, editors, *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World* (Basingstoke: MacMillan, 1986); Bruno Latour, *Science in Action: How to Follow Scientists and Engineers Through Society* (Cambridge, MA: Harvard University Press, 1987); Strum and Latour; Latour, 'Ethnography of a High-Tech Case'; Bruno Latour, 'On Recalling ANT', in: John Law and John Hassard, editors, *Actor Network Theory and After* (Oxford: Blackwell, 1999); and Latour, *Reassembling the Social*. After criticizing every element of the name, including the hyphen, in 1999, Latour now ardently defends it, even to the point of adding yet another hyphen between the last two words.

"speaks for," the actors, providing them with new roles and new sets of interests. Similar to the hierarchical nature of systems, actors-networks (e.g., a corporation or a computer) may themselves be actors in another actor-network (e.g., an industry or a transaction-switching network).

Actors may of course resist their enrollment, and thus alter or completely destroy the nascent network. This is readily understandable when speaking of animate actors, but ANT's methodological parallelism requires the same language to be used when speaking about material artifacts, such as batteries, "resisting" their assigned role as an adequate power source for an electrically powered vehicle.[32] This may be a poetic way to refer to the obduracy of the material world, but it also has the problematic side-effect of assigning agency to inanimate objects.

Although the concepts of actor-networks and sociotechnical systems are very similar, Callon argues that the former avoids the methodological problem of having a context or environment that affects the system, but is not part of it.[33] Hughes's response, already described above, was to distinguish between interaction and control, noting that while certain elements will necessarily remain outside the direct control of the system builders, they nevertheless affect the system. These elements can thus be taken as "given" in some sense by the historian, as the system builders have (or assume they have) no ability to change them.[34] However, historians should always keep in mind that what is context for some might be content for others. Making one's focus explicit largely resolves this issue.

Law also preferred the concept of an actor-network to a system because the former emphasizes the difficulty network builders face trying to translate the various actors into their respective roles.[35] Law claimed that systems are often assumed to be self-organizing and self-sustaining, while actor-networks are inherently fragile. Hughes's concept of technological momentum is likely to blame for this impression. Hughes argued that when a sociotechnical system matures, groups often develop vested interests in its preservation, creating a kind of "mass" that ensures a self-sustaining "inertia of directed motion."[36] This phenomenon is most certainly observable, but General System Theory does not assert that self-maintenance is a general quality of *all* systems.

2.2 Methodology

The approaches described in the previous section are useful for analyzing information about a historical case, but only if that information was gathered properly. Thus, we must now turn to

[32] Callon, 'The Sociology of an Actor-Network'.

[33] Michel Callon, 'Society in the Making: The Study of Technology as a Tool for Sociological Analysis', in: Weibe E Bijker, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987).

[34] Hughes, 'Seamless Web'.

[35] Law.

[36] Hughes, *Networks of Power*, p. 15.

the specific sources and methods I used to gather and analyze information about Visa's history, the known problems with these, and the ways in which I sought to overcome those problems as much as possible.

The primary sources for this project can be grouped roughly into three types: news and trade publications from the period; interviews with those who designed, built, and participated in the Visa payment system; and a selection of primary documents that I was able to obtain. Each of these types offer specific opportunities as well as dangerous pitfalls that I will discuss in turn. In the final section I will describe the techniques I used to put these three sources together into a cohesive narrative.

2.2.1 News and Trade Publications

The first set of sources I consulted were the key news and trade publications from the time period covered by this dissertation (1970–1984). The most useful by far was the *American Banker*, a daily newspaper covering the US banking industry, which is currently indexed in printed form back to 1971. Another source commonly used by other authors is *The Nilson Report*, a twice-monthly newsletter on the payment card industry published since 1970. Unfortunately, the only publicly available archive of this newsletter is limited to 1976 through 1984, and I had no budget with which to purchase the earlier issues.[37] However, the missing period was well covered by the *American Banker*, as well as the *Payment Systems Newsletter*, a now-defunct publication that competed with *The Nilson Report* at the time. Various publications from the American Bankers Association (ABA), such as handbooks and training manuals, published during this period also provided brief histories as well as insights into the organizational, economic, and legal issues that concerned bankcard program managers at the time.[38]

Although these sources discussed events from the banker's perspective, the merchants and the cardholding public often held a different view. To understand the perspectives of these other groups, I consulted publications such as *Business Week*, *Forbes*, *Life*, *Time*, and *Newsweek*. Stallwitz's survey of San Francisco Bay Area merchants in 1968 provided further insights not only into their reasons for participating or abstaining at that time, but also their general attitudes towards the payment card networks.[39] The perspectives of merchant and consumer affairs lobbying groups can also be seen in various Congressional hearings, but one should not assume that these groups speak for all merchants and consumers.[40]

[37] Although Nilson had previously provided access to Mandell and Nocera, the current management no longer grants free access to historians.

[38] For example, see Lewis Mandell and Neil B Murphy, *Bank Cards* (Washington DC: American Institute of Banking and American Bankers Association, 1976); John Fisher et al., *Bank Cards* (Washington DC: American Bankers Association, 1980); and *Critical Issues in Bank Cards* (Washington DC: American Bankers Association, 1980).

[39] John F Stallwitz, 'A Market Research Study of the Attitudes and Reactions of Member and Non-Member Merchants to Bank Credit Card Plans', Masters thesis, University of California Berkeley (December 1968).

[40] *Unsolicited Bank Credit Cards*, Hearings before the Committee on Banking and Currency, United States House

The news and trade publications were useful for a number of reasons. First, they helped establish a basic timeline of events that could be used to guide the interviews, and correlate the stories told by the informants. Second, they brought forth the various issues and debates that concerned the actors *at the time*, which were often forgotten by those same actors when interviewed thirty-five years later. Third, they also made apparent the exact language and categories actors used during those debates, some of which shifted or mutated over the years. The language used by the actors at the time often revealed much about how they understood the contested issues, or the very nature of the payment card systems themselves.

This is not to say, however, that the news and trade publications should be treated as impartial or infallible. Reporters and their publications have interests and agenda just as the direct participants do. They may show bias towards a particular group, or refrain from reporting on certain topics, in order to appease or placate their advertisers. Conversely, they may reprint press releases from organizations without critically evaluating them, lending a degree of impartiality to what is essentially a marketing statement. Reporters unfamiliar with the new technologies being developed may also make mistakes when describing them, some of which may never be retracted. Even worse, future articles that draw upon the original may continuously repeat the error, giving it the appearance of an established fact.[41]

Compensating for these dangers is difficult, but not impossible. Understanding the publication's interests, along with careful and close reading, can reveal biases, at least to some degree. Factual statements within a story can be checked for internal as well as external consistency. Clumsy or inappropriate language can indicate that the reporter does not fully understand the subject matter. Direct repetition of phrases can often signal reliance upon previous articles. Ultimately, information gathered from news sources must be treated as one account of many, and the techniques I used to combine these sources with the other types will be discussed in the final section.

2.2.2 Autobiographies and Oral Histories

Another group of sources I gathered and analyzed for this project consists of three related but distinct elements: the autobiography of Visa's founder and first CEO, Dee Hock; a series of stories about Visa's early years written by their treasurer and CFO, Tom Cleveland; and numerous interviews I conducted with those who designed, built, operated, and participated in the Visa payment system. I discuss them together in this section because they all share a common set of opportunities and dangers.

Hock's autobiography was first published in 1999, and was updated and republished in

of Representatives, 19th Congress, First Session. (8 and 9 November 1967); *The Consumer Credit Protection Act Ammendments of 1977*, Hearings on HR 8753 before the Subcommittee on Consumer Affairs of the Committee on Banking, Finance, and Urban Affairs, United States House of Representatives, 95th Congress, First Session. (March 1977).

[41] One example of this appears in section 3.2.3. The often-reported inspiration for the Diners Club was actually concocted by their press agent.

2005 under a different title.[42] It is a key source for a number of reasons. First, it is the only source that discusses his early life and the episodes that led him to adopt various convictions about organizations and the nature of money that would eventually shape the design of Visa. Second, it reveals how Hock understood the problems facing the BankAmericard system in 1968, which helps explain why he proposed the radical solution of a new organization. Third, it explains his intentions behind the structure of Visa and what he was ultimately trying to accomplish, which as we shall see, was far larger than most of the member banks assumed.

Cleveland's stories were written in 1999, just before his retirement from Visa, and they attempt to convey "the humor and essence of the 'Dee Hock' experience."[43] They are not strictly autobiographical, but they are written in retrospect, and thus face some of the same issues as autobiographies and oral histories. Cleveland admits that they are written entirely from his own perspective, but that perspective is rather valuable, as it reveals much more of the gritty political and cultural dynamics than Hock's autobiography. Not surprisingly, his portrayal of Hock also differs markedly from Hock's own autobiographical persona. Some of the stories are shockingly humorous, but they help to lend a human face to an organization that is often portrayed as being above such things. Lastly, they also provide an internal financial perspective on some key events during this period that is missing from other sources.

My own interviews complemented these other two sources by adding the perspectives of 45 other informants, many of whom were interviewed several times.[44] Over the course of two and a half years, I conducted a total of 60 interviews, most of which lasted for an hour and a half, but several lasted multiple hours, the longest of which consumed an entire day. Most informants were previous employees of Visa, though a few were involved with card programs at a member bank.[45] Most were key decision makers in the organization, though some were rank-and-file computer programmers or operators. Most were men, but perhaps surprisingly for the time, six were women.

These autobiographies and oral histories enhance the news accounts in several ways. First, they add significant context and depth to the reported details. They reveal new information about *why* certain decisions were made, why other paths were not followed, and how certain objectives were actually accomplished. By speaking with the pioneers, I was able to learn more about their original intentions and motivations, as well as the meaning they now, in retrospect, attach to their actions. They are understandably proud of what they accomplished, and took great joy in explaining the finer and more subtle aspects of the payments industry at that time.

Second, and perhaps most important, the oral histories reveal many details that were simply

[42] Dee Hock, *Birth of the Chaordic Age* (San Francisco: Berrett-Koehler, 1999); Hock, *One From Many*. The second edition has a slightly stronger emphasis on his "chaordic" concepts, which will be discussed in Chapter 4.

[43] Tom Cleveland, 'The Visa History: Tom Cleveland's Perspective' (April 1999).

[44] Although I was able to interview Cleveland, Hock declined to be interviewed, referring me to his autobiography.

[45] For details, see the list of interviewees at the end of the dissertation. The member banker perspective is somewhat underdeveloped in this dissertation. In future research, I hope to interview more of them to learn how they viewed the events during this period.

never publicly reported, and many times, never written down at all. In the early years, Visa was a small, highly-entrepreneurial organization in a rapidly developing industry, and employees rarely had time to document their decisions or actions, much less save them in a systematic archive.[46] Often, the only place these details still exist are in the stories told by the actors themselves.

Third, the oral histories provide a different, less-guarded perspective on the events of this period. None of the informants still work for Visa, and most had long-since retired by the time of the interview. Many were willing to speak frankly about what they did and why they did it, as they no longer felt a need to represent the interests of the organization, nor feared any reprisals. Most were exceptionally candid about Hock's personality and behavior, which provided a much-needed balance to his own autobiography.

Although autobiographies and oral histories can significantly enhance the details one can garner from news accounts, they also bring with them some significant dangers.[47] First, and most obvious, is the problem of faulty and reconstructed memories. Portelli reminds us that the mind is not like a video recorder: "memory is not a passive depository of facts, but an active process of creation of meanings."[48] In order for something to stay in our memory, it must "mean" something significant to us—that is, it must help us make sense of our lives.[49] Thus, our reconstructed memories of past events will necessarily differ from those of other participants. Smith and Watson build upon this, stressing that autobiographies "offer subjective 'truth' rather than 'fact.'"[50] Lummis further warns that we continue to alter or enhance our memories as they are recalled over time, refitting them into our growing and shifting identities.[51] Portelli concludes that oral histories "tell us less about *events* than about their *meaning*...[they] tell us not just what people did, but what they wanted to do, what they believed they wanted to, what they believed they were doing and what they now think they did."[52]

Our individual meaning-making may be further influenced by the social and political structures that surround us, both at the time of the events and at the time of recall. Bodnar found that his informants "did not remember alone," observing that their memories tended to align with

[46] Visa is also reticent to share whatever internal documents do remain from that time, as they are currently restructuring into a publicly-held firm, and are thus understandably wanting to control their image.

[47] For an introduction to the issues of autobiography, see Sidonie Smith and Julia Watson, editors, *Getting a Life: Everyday Uses of Autobiography* (Minneapolis, MN: University of Minnesota Press, 1996); and Sidonie Smith and Julia Watson, *Reading Autobiography: A Guide for Interpreting Life Narratives* (Minneapolis, MN: University of Minnesota Press, 2001), especially Chapter 7, which provides a "toolkit" for reading life narratives. For an overview of oral historiography, see Paul Thompson, *The Voices of the Past: Oral History*, 3rd edition (Oxford: Oxford University Press, 2000); Robert Perks and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998); and Jan Vansina, *Oral Tradition: A Study in Historical Methodology*, Translated by H M Wright (London: Routledge, 1965).

[48] Alessandro Portelli, 'What Makes Oral History Different?', in: Robert Perks and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998), p. 69.

[49] Karl E Weick, *Sensemaking in Organizations* (Thousand Oaks, CA: SAGE, 1995).

[50] Smith and Watson, *Reading Autobiography*, p. 10.

[51] Trevor Lummis, 'Structure and Validity in Oral Evidence', in: Robert Perks and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998).

[52] Portelli, p. 67.

the rhetoric of the institutions to which they belonged, but became more fragmented when the power of those institutions waned.[53] Thomson showed how powerful national and social myths caused ANZAC soldiers to reconstruct or suppress their memories of World War I, even to the point of describing scenes from the movie *Gallipoli* in the first person.[54] Our memories may be further influenced by institutional and organizational myths that seek to give their members a sense of identity and purpose.[55]

The problems of retrospective narration are legion, but this is not to say that what we get from autobiography and oral history is somehow useless or irrelevant. As with any source, the content must be checked for internal and external consistency, but the degree to which a retrospective account differs from earlier ones can also be enlightening. The difference may be due to justifications, rationalizations, or a desire to suppress unpleasant memories. However, Thompson argues that informants may actually develop a more clear and balanced picture of past events as they are further removed from the social influences of the time.[56]

A related but more subtle danger posed by autobiographies and oral histories is the problem of self-revelation. If we look to these sources to learn not only about events, but also about an informant's personality or character, we must be aware that what is revealed often tells us more about who the informant would *like to be*, rather than who he or she really *was or is*. Swindells reminds us that autobiography is "something more than a simple presentation of individual existence."[57] As with many social interactions, the author puts forth a persona that the intended audience would find not only understandable but also attractive. As Smith and Watson argued, "In telling their stories, narrators take up models of identity that are culturally available."[58] Authors put on these borrowed identities in order to portray themselves as attractive archetypes rather than as the messy and complex characters they actually are.

A final danger, which is unique to oral histories, is the historian's participation in the interview process. Oral histories are more than simply autobiographies expressed in a different medium—they are "collaborative histories" jointly created by the informant and the historian.[59] As a result, there are several ways in which the historian can unwittingly influence the content of the informant's memories during the interview, making it less useful for future analysis. First, the location and social context of the interview may alter what the informant is willing to share.[60] Second, the questions asked by the historian obviously influence the reply, but they may also create confusion if they are phrased in the historian's categories instead of the in-

[53] John Bodnar, 'Power and Memory in Oral History: Workers and Managers at Studebaker', *Journal of American History* 75:4 (March 1989).

[54] Alistair Thomson, 'Anzac Memories: Putting Popular Memory Theory Into Practice in Australia', in: Robert Perks and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998).

[55] For the role of "saga" in organizational identity, see Clark. For the ways in which workers resist these identities, see Martin Parker, *Organizational Culture and Identity: Unity and Division at Work* (London: SAGE, 2000).

[56] Thompson.

[57] Julia Swindells, editor, *The Uses of Autobiography* (London: Taylor and Francis, 1995), p. 2.

[58] Smith and Watson, *Getting a Life*, p. 9.

[59] Thompson.

[60] Ibid., pp. 125–128.

formant's.[61] Third, the interview setting may create "unspoken assumptions" about what the historian wants to hear, encouraging informants to tell certain stories, even if they are exaggerated or based on rumor.[62] Lastly, although showing empathy towards the informant may build trust and thus encourage deeper sharing, it may also cause the historian to be "captured" by the informant, and drawn into taking a particular side in a historical controversy.[63]

Although I did observe many of these dangers while conducting and analyzing the interviews, they were actually less pronounced that I had expected. As noted earlier, most informants had been retired for at least a few years, so most were quite comfortable with discussing details and admitting mistakes. The interview locations were chosen by the informants, but since none still worked for Visa, they were not constrained by concerns over reprisals. After a few initial interviews, I became proficient in "Visa speak" and was typically able to frame questions in the informant's categories. I designed my opening questions so that the informant could develop their own themes, and when I did have to ask more leading questions to get to the heart of an issue, the informants were always quick to correct any misinformed assumptions. My greatest danger was being captured by the informants, as I respect them greatly, but I tried to balance what they said against other testimonies, what I read in the news sources, and the primary documents I was able to obtain.

Faulty and reconstructed memories were certainly apparent, but informant's memories were notably consistent with each other, as well as the written record, when it came to "firsts." For example, memories about the first computerized authorization system were highly correlated with the surprisingly detailed press reports at the time.[64] The same was true for the first computerized clearing and settlement system.[65] However, memories became less consistent, and the order of events became more hazy, when it came to more incremental changes, such as the development of point of sale dial terminals.[66] Obviously, some informants remembered more clearly than others, and it was relatively easy to discern during the interview how sure the informant was about the content of their stories from their pauses, phrasing, volume, pitch, and body language.

Although their memories were no doubt shaped to some degree by the social and political structures that surrounded them, some were surprisingly eager to dispel certain myths about Visa and its founder that have been perpetuated in other sources. Few felt any need to protect the interests of the current organization, nor those who once held power over them, but they

[61] Portelli. See also Friedlander, 'Theory, Method and Oral History' and Grele, 'Movement Without Aim' in the same volume.

[62] Thompson.

[63] For example, see the debate between Martin and Collins: Brian Martin, 'The Critique of Science Becomes Academic', *Science Technology and Human Values* 18:2 (1994); H M Collins, 'In Praise of Futile Gestures: How Scientific is the Sociology of Scientific Knowledge?', *Social Studies of Science* 26 (1996); Brian Martin, 'Sticking a Needle into Science: The Case of Polio Vaccines and the Origin of AIDS', *Social Studies of Science* 26 (1996).

[64] See Chapter 5.

[65] See Chapter 6.

[66] See Chapter 8.

were never vindictive. As would be expected, they did want to dwell on their successes more than their failures, but all were honest about their shortcomings when asked. Lastly, key informants reviewed drafts of the following chapters to ensure that I had not added erroneous layers of meaning, and that their comments were represented fairly.

2.2.3 Primary Documents

The last set of sources were a selection of primary documents created by the key actors near the time of the events. These included reprints of speeches given by Hock and other executives, project reports, statistics, brochures produced for the member banks, consumer survey results, as well as artifacts such as old cards and sales drafts. These sources enabled me not only to corroborate statements made by the interviewees, but also to discover new, detailed information that did not appear elsewhere.

The speeches in particular were especially useful for validating claims made by Hock in his autobiography.[67] His radical ideas on the nature of money and organizations are actually evident in his speeches from the 1970s, although the ideas were often less clearly-articulated then. Thus we can have some assurance that these ideas played a significant role in his decision making, and were not simply reconstructions of, or projections onto, past events.

2.2.4 Techniques for Building the Narrative

After gathering and analyzing these sources, I began to fit them together into my own narrative. I started with the basic chronology of events gleaned from the press reports, establishing a rough timeline. Other events revealed by the autobiographical, oral, and primary document sources were then added. Wherever possible, exact dates, statistics, or factual claims came from printed sources created near the time of the events. I then grouped these events into topic areas, which were assigned to chapters based on common themes.

While writing each chapter, I consulted all the relevant sources and asked numerous followup questions of the appropriate informants to triangulate as accurate an account as possible. Disagreement between sources was surprisingly infrequent, but when it did occur, I assigned relative weighting according to a number of factors: general reliability of the source; how close the source was to the events in question; how likely it would be for the source to know the information they claimed to know; the degree to which the source seemed sure of their story during the interview; etc. When determining a single account proved impossible, I noted the various opinions, and left the question open.

Although the methods discussed in this section should, I hope, lead to a narrative that is as accurate as possible, no method will produce a true and accurate description of the time period "as it really was." This is widely recognized as a chimera in historiography, and I

[67] Hock, *One From Many*.

only claim that what follows is *a* history of the Visa payment system, not *the* history.[68] Like any history, this history is told from my own perspective, which is highly influenced by my academic training in Science and Technology Studies, and my technical training in information systems and software engineering.

2.3 Conclusion

In this chapter, I showed that there is not yet a history of the Visa payment system told from the STS perspective, and through my review of the key STS literature on technological change, I sketched the STS perspectives that are most relevant to the case. I then described the sources used for this project, and my methods for gathering and analyzing them.

In the next chapter, we will augment this theoretical context with a historical one. Specifically, I will review the important innovations that occurred within the American payments industry directly leading up to the formation of the organization now known as Visa.

[68] E H Carr, *What is History?* (New York: Penguin Books, 1961).

Chapter 3

Historical Context: The American Payments Industry Before Visa

In the previous chapter, I began to construct a context from which we can understand the detailed history presented in Part II. The theoretical portion of that context consisted of two key approaches from the Science and Technology Studies literature that are helpful for understanding the origins and evolution of sociotechnical systems in general. In this chapter, we now turn our attention to the specific kind of sociotechnical system Visa built: a payment card system. In order to understand their system, and why they built it the way they did, one first needs to understand a bit about payment card systems in general, and how they developed in America throughout the twentieth century.

This chapter is structured into two main parts. First, I will introduce the basic terms and concepts used to describe and analyze payment card systems. Second, I will review the important innovations that occurred within the American payments industry directly leading up to the formation of the organization now known as Visa. These innovations formed a historical context that influenced not only the way Visa was formed, but also the way it continued to develop.[1]

3.1 Basic Terms and Concepts

The payments industry, like most others, uses a complex set of terms whose meanings may not be obvious to outsiders. This section will review the most basic of these terms that are common to nearly all payment card systems. Other terms that are more specific, or best understood within their historical context, will be introduced in the next section.

The best place to start is with the term *payment system* itself. Simply put, a payment system

[1] The material in this chapter comes mostly from secondary sources, though I will often augment and critique these through the use of primary sources I uncovered during my research.

is a set of social relations and technological mechanisms that enable the exchange of value.[2] Thus payment systems are sociotechnical systems (i.e., involving both social and technical elements) designed to serve a particular purpose: the exchange of value between parties. The "social relations" within a payment system consist of the rules that govern how the various parties and their agents are organized and represented, how decisions are made, how exchanges take place and how grievances are settled. These rules may be established informally through a culture (as in the case of barter and early currency systems), or they may be formalized into laws, contracts, or other legal agreements (as in the case of cheque or payment card systems).[3] The "technological mechanisms" within a payment system not only enable the value exchanges to take place, but also tend to make them more efficient. These might be relatively simple technologies such as commodity money, coins or notes; or they might be complex technological systems involving powerful computers and large telecommunication networks.[4]

A *payment card system* then is a type of payment system where a device such as a card is used to identify uniquely a buyer's account to the system.[5] Although these devices have typically been fashioned as thin rectangular cards similar to those exchanged by business people, they need not be of any particular form. Any device that can communicate the buyer's account information could serve the same function. As we shall see, payment "cards" throughout history have actually taken many forms: paper cards, metal fobs, metal address plates, embossed plastic cards, and plastic cards with embedded chips. Some organizations are now experimenting with the use of mobile phones, radio transmitters used to pay road tolls, and tiny RFID chips embedded in wrist bands or even under the skin.[6]

[2]This definition comes partly from Humphrey, partly from Hock, *One From Many*, and partly from interviews with Tom Honey and other Visa pioneers.

[3]Of course, many systems have both—some rules are codified and formalized while others are inherent in the culture.

[4]A true philosophical inquiry into money, and a detailed history of its development, is unfortunately beyond the scope of this dissertation. For a philosophical treatise on money and its relationship to value, see Georg Simmel, *The Philosophy of Money*, Third Enlarged edition (1978; reprint, London: Routledge, 2004). For a sociological look at how money is used, see Viviana A Zelizer, *The Social Meaning of Money: Pin Money, Paychecks, Poor Relief and Other Currencies* (New York: Basic Books, 1994). For a philosophical and literary review of the meaning of money, see James Buchan, *Frozen Desire: The Meaning of Money* (New York: Farrar, Straus and Giroux, 1997). For theoretical texts on money and banking, see W T Newlyn, *Theory of Money*, 2nd edition (1962; reprint, Oxford: Clarendon, 1971); Richard Coghlan, *The Theory of Money and Finance* (London: Macmillan, 1980); Lester V Chandler, *The Economics of Money and Banking*, Ninth edition (1948; reprint, New York: Harper and Row, 1986); and Raymond P Kent, *Money and Banking*, 4th edition (1947; reprint, New York: Holt, Rinehart and Winston, 1961). The most complete and authoritative review of the history and anthropology of "primitive money" is Paul Einzig, *Primitive Money in its Ethnological, Historical and Economic Aspects*, 2nd edition (1949; reprint, Oxford: Pergamon, 1966). For other histories of money, see Glyn Davies, *A History of Money: From Ancient Times to the Present Day* (Cardiff: University of Wales Press, 1994); John Chown, *A History of Money: From AD 800* (London: Routledge, 1994); John Kenneth Galbraith, *Money: Whence it Came, Where it Went*, Revised edition (1975; reprint, Boston: Houghton Mifflin, 1995); and Dennis W Richardson, *Electric Money: Evolution of an Electronic Funds-Transfer System* (Cambridge, MA: MIT Press, 1970). For an intriguing argument that money developed out of religious ritual and not economic necessity, see William H Desmonde, *Magic, Myth, and Money: The Origin of Money in Religious Ritual* (New York: Free Press of Glencoe, 1962).

[5]Note that banks often refer to their card systems as "card programs." When discussing bank payment card systems, I will occasionally use this term.

[6]The use of toll radio transmitters is discussed in Andrei Hagiu, 'Cash, Credit, or Car?' ⟨URL: `http://www.`

Although a payment card system is similar to a cheque system, it is important to note that a payment card is not itself a payment instrument in the same way a cheque is. A cheque both identifies the customer and acts as the payment instrument. In a payment card system however, identification is separated from the payment instrument. Thus, a payment card's primary function is to establish account *identity*, which is then recorded onto a separate form that becomes the actual payment instrument. Although it could not have been anticipated, this separation was necessary to move from paper-based transactions to fully-electronic ones.[7]

Throughout the history of these payment cards, a variety of terms have been used to describe them. Often these delineate the cards based on what kinds of funds they access: *charge cards* are those that access a short-term credit account that must be paid off at the end of each billing cycle; *credit cards* are those that offer a long-term revolving credit line, requiring only a minimum payment each month; *debit cards* are those that access funds in deposit or semi-liquid investment accounts;[8] and *stored-value cards*, sometimes called *cash cards*, access funds that have been in effect transferred to a special account accessible only by that card.[9] Although we typically use these delineating terms today with relative precision, historically they were often muddled together because the various actor groups had not yet created these distinct categories. For example, early travel and entertainment cards were often called "credit cards" because they offered short-term credit; only after the creation of revolving credit cards was the term "charge card" used. In order to avoid confusion, I will use these distinctive terms when discussing various historical payment cards, except where it would be historically misleading to do so, such as when discussing the debates surrounding Visa's proposed "asset card." However, it should be noted that actors may not have used those distinctive terms at the time, and one must be careful when reading early sources that use the term "credit card."

Payment cards are distributed to consumers, referred to as *cardholders*, by organizations

`marketplatforms.com/payingwithplastic/mpdperspectives/`⟩ – visited on 7 September 2006. Resorts such as the Great Wolf Lodge at Niagara Falls, Canada, are now offering wrist bands with RFID transmitters to their guests that serve not only as their room key, but also as their identification when paying for goods and services within the resort ('Great Wolf Lodge Rides the RFID Wristband Wave' ⟨URL: `http://www.pdcorp.com/crowd-control/case-study-great-wolf.html`⟩ – visited on 13 September 2006). Several Europeans clubs and resorts are going a step further, offering to implant an RFID chip contained in glass directly under the skin, enabling customers to charge goods and services by simply walking up to a scanner (Andrew Losowsky, 'I've Got You Under My Skin', *The Guardian* (10 June 2004), ⟨URL: `http://technology.guardian.co.uk/online/story/0,3605,1234827,00.html`⟩ – visited on 13 September 2006).

[7] Note that the US government recently passed laws that allow merchants to scan cheques at the point of sale and process the transaction completely electronically. The cheque in this scenario merely becomes an identity device.

[8] Note that there is great variability as to *when* funds are actually removed in response to a debit transaction. Debit transactions may be processed offline or online, and local banking laws may also restrict how quickly funds may be removed from an account.

[9] Very few stored-value cards actually store the available funds on the card itself. It is much more common that the card accesses a special account, so that loosing the card does not necessarily result in loosing funds. The key distinction between debit and stored-value cards is that the stored value account can be accessed only by the stored value card, where the debit account may be accessed by other payment devices, such as cheques (Elliott interview). The various terms for payment cards come from Evans and Schmalensee. Visa also developed travel money and gift cards, which function essentially like stored value cards, accessing special funds that have already been deposited into a private account.

called *issuers*. Cards have been issued by a great variety of organizations throughout history: banks and other financial firms, retailers, oil companies, transportation companies, and telecommunication companies, just to name a few. Any organization that develops a financial relationship with consumers could issue a card for use in an established payment system, or in a new one they create.[10] Although there are only a few national or international payment card networks, numerous localized, small-scale networks are constantly being developed, such as those used within academic institutions, resorts, theme parks, and transportation networks.[11]

Cardholders use their cards at *merchants*, which is the generic term used in the payments industry for someone on the receiving side of a transaction (i.e., someone who is being paid). The merchant is then represented in the payment system by an organization called the *acquirer* or *merchant bank*.[12] The acquirer maintains the merchant's account, crediting or debiting it for transactions processed through the system, minus any fees. However, the actual processing of the transaction may be done by a third-party organization called a *processor*, who may serve the merchant acquirer, the issuer or both. Processors specialize in high-volume transaction processing and thus offer an attractive level of expertise and economy of scale to smaller issuers and acquirers. As payment card systems became electronic in the 1970s and 1980s, banks increasingly outsourced their processing to third parties, primarily because that activity began to resemble computer data processing more than traditional banking operations.

While the terms issuer and acquirer often imply that these are two different organizations, they need not be. Issuer and acquirer are actually *roles* played by an organization in the context of a particular transaction. The same organization may be both the issuer and the acquirer, or may be an issuer in one transaction and an acquirer in another. For example, American Express is both the issuer and acquirer for every transaction in their system, and Bank of America both issues cards and acquires merchant transactions in the Visa system. One of the central principles of the bankcard systems is that all member organizations participate in, or at least support, both the issuing and acquiring sides of the business. However, one of the trends in the Visa system has been the specialization of members into being solely issuers or solely acquirers, the reasons for which will be discussed in Chapter 8.

3.2 Innovations in the American Payments Industry, 1913–1969

Now that we have a basic understanding of the terms used in payment card systems in general, we now turn our attention to the historical context surrounding the creation of the Visa system in particular. This section will review the various innovations that took place within the Amer-

[10] Note that some established payment card systems like Visa have restrictions on who can be an issuer.

[11] For a history of how the Octopus transportation system's card began to creep out to nearby merchants, see Lucia Siu, 'Octopus and Mondex: The Social Shaping of Money, Technology and Consensus' (April 2002), ⟨URL: `http://www.sociology.ed.ac.uk/finance/Papers/Siu_Octopus.pdf`⟩ – visited on 4 September 2006.

[12] Although the concept of the acquirer also exists within cheque payment systems, we will concentrate on their role payment card systems in this section.

ican payments industry starting with the creation of the Federal Reserve's centralized, national cheque clearing system. The following five decades saw an unprecedented explosion of innovation in the payments industry, from the early retail charge cards, to the more general-purpose travel and entertainment cards, to bank issued credit cards featuring revolving credit lines, to inter-bank associations that extended the reach of bankcards across the nation and beyond. This section will review these various innovations in detail so that we can better understand the historical and competitive context surrounding the creation of Visa and its payment system.

3.2.1 The Federal Reserve System and Cheque Clearing

The payments industry, as well as the entire banking system in America, was radically altered by the formation of the Federal Reserve System in 1913. The Federal Reserve (the Fed) changed not only the way funds were transfered between banks, but also the way cheques were cleared and settled, creating the economic conditions that would subsequently make cheques the predominant method of payment in America throughout the rest of the twentieth century, and eventually a thorn in the side of debit card transaction pricing. However, to understand this development fully, we first need to understand a bit about how cheques operated in the American payment system before the creation of the Fed.

A cheque is a different kind of payment instrument than cash. A cheque is actually a modern form of the medieval "bill of exchange," first developed in twelfth-century Italy.[13] A cheque is an instruction, to the bank upon which it is drawn, to pay the bearer the specified amount upon presentment. As opposed to cash, a cheque does not directly represent value. Rather it is a *tentative claim* on value and in order for the holder of a cheque to receive actual funds, the cheque must go through a clearing and settlement process.

Clearing a cheque means routing it to the bank upon which it is drawn in order for that bank to ascertain if the cheque is valid and if the account has sufficient funds for payment. This sounds simple enough, but cheque clearing in the nineteenth and early twentieth centuries was an amazingly complicated process due to the laws that governed how cheques were paid. According to those laws, in order to receive the full value written on the cheque, the holder (or an agent of the holder) had to present the physical piece of paper to the bank upon which it was drawn. The paying bank was then required to clear the cheque *at par*, meaning that they paid the full value of the cheque to the bearer. If the cheque was presented through the mail or by a third-party courier, the paying bank was allowed to *discount* the cheque, meaning that it would pay something less than the full face value. This discount ostensibly helped cover the costs and risks to the paying bank but it was also an attractive source of additional income. The presenting bank would typically pass on these discount fees to the depositor, which might be a significant portion of the cheque's value. Thus there was little incentive for payees to accept cheques drawn on banks other than their own, unless their bank had a way to avoid paying the

[13] Davies, p. 156; Evans and Schmalensee, p. 28.

discounts.[14]

One way in which banks avoided paying discounts on local cheques was to send messengers to each bank each day to present their cheques in person. In areas where multiple banks operated, the banks typically formed cooperative *clearinghouses* in order to make the process more efficient. Instead of each bank sending messengers to every other bank, all the messengers assembled at one location and simply exchanged each other's cheques.[15] Although the messengers were no longer technically presenting the cheques at the bank itself, the banks saw no value in charging discounts on each other, as they would only tend to cancel each other out. Thus it was common that cheques presented to clearinghouses were cleared at par.

The way in which banks avoided paying discounts on out-of-town cheques has more to do with the way these cheques were *settled*, which is the process of transferring "good and final funds" to the presenter.[16] Instead of shipping gold or notes around the country, most banks by the early 1900s had established *correspondent relationships* with banks in other cities in order simplify the process of long distance funds transfer. Correspondent banks make deposits with each other, and then debit or credit those deposits when they need to transfer funds. For example, a bank in Boston might deposit funds with a bank in New York, and vice versa. When the Boston bank needed to transfer funds to the New York bank (perhaps to settle a cheque), the New York bank would simply debit the Boston bank's account and credit the the account of the payee. Thus, transferring funds became a simple bookkeeping entry instead of a physical movement of currency. Because the transfer was relatively easy, and because correspondent banks wanted to maintain a congenial relationship, they would settle each other's cheques at par, even if they were sent via a third-party courier.

The combination of correspondent relationships and clearinghouses created a complex, web-like network of banks willing to clear cheques at par, and not surprisingly, banks went to great strides to leverage this network to avoid incurring discounts. There are legendary stories from the time of cheques traveling ridiculous distances over circuitous routes to get to a paying bank that was relatively close to the originating bank. One story described a cheque that had travelled 1,500 miles over 11 days to get to a paying bank that was only 100 miles away.[17] Another story told of a cheque that travelled 4,500 miles over two weeks to get to a competing bank that was only 4 miles away, only to find that there were insufficient funds, resulting in its return via the same route.[18] These delays were actually welcomed by the cheque payors and

[14] Evans and Schmalensee, p. 38.

[15] This was by no means a new idea. The Bankers' Clearing House in London began in the 1770s. See Martin Campbell-Kelly and William Aspray, *Computer: A History of the Information Machine* (New York: Basic Books, 1996), pp. 15–18.

[16] Humphrey, p. 4.

[17] Evans and Schmalensee, pp. 40–41.

[18] Benjamin J Klebaner, *Commercial Banking in the United States: A History* (Hinsdale, IL: Dryden Press, 1974). See also Leonard W Fernelius and David Fettig, 'The Dichotomy Becomes Reality: Ten Years of the Federal Reserve as Regulator and Competitor' ⟨URL: `http://minneapolisfed.org/pubs/ar/ar1991.cfm`⟩ – visited on 26 August 2006.

their banks, as it allowed them to hold onto the funds for a longer period of time. This delay between the time a cheque is deposited and the time funds are finally settled is called *float*, and the benefits payors realized from an elongated float period could add up to a substantial amount.

The Federal Reserve System was established at the end of 1913 in an attempt to reform both the US banking system and the way cheques were cleared. The Fed altered the nature of American banking in many ways, but they had two important effects that are relevant to payments and cheque clearing. First, the new Fed structure created a mechanism for funds transfer that was as easy and efficient as the correspondent model, but with a simple hierarchical organization and a national scope. All nationally-chartered banks were required to become members of the Federal Reserve System by maintaining a reserve deposit account at their local Federal Reserve District Bank.[19] These reserve accounts created a way for the Fed to not only control the money supply, but also transfer funds between member banks in the district by simply debiting one reserve account and crediting another. The District Banks then maintained reserve accounts with each other, which allowed for easy transfer of funds between banks in different districts. This system almost completely eliminated the need to move physical currency around the nation.[20]

This new funds transfer system had a subtle, though perhaps unintended effect on the nature of money: it made it more abstract. Within the Federal Reserve System, money was now less of a "thing" and more of a concept.[21] The transfer of money no longer required the movement of physical objects, only the mathematical manipulation of numbers written in an account book. For the average consumer, "money" was still physical notes and coins, but for those involved in the operation of the banking system, money was quickly becoming something more akin to socially-guaranteed information.[22] The processing of payments using computer networks depends on just this kind of conceptual move, and as we shall see, creating a similar conceptual move within the minds of consumers was one of Visa's central missions. However, this is not to say that the Fed made this move with electronic payments in mind, or that we should interpret this move as being the first step on an inevitable path towards electronic value exchange. Rather, we should see it as a conceptual move that occurred for reasons specific to the time, but was nevertheless a necessary condition for future events that took place in the history of Visa and other electronic payment systems.

The second effect the Fed had on payments was the establishment of a centralized, national cheque clearing system in 1915 that promised to be more efficient than the haphazard method

[19] State-chartered banks were also encouraged to join, but most did not as the reserve requirements for Federal Reserve members were often higher than their state-mandated counterparts.

[20] Evans and Schmalensee, pp. 38–40. See also, Humphrey, p. 8.

[21] A number of authors on the history of money remark that this move towards abstraction has been a general trend. See especially Richardson.

[22] This is also the conclusion reached by Dee Hock, Visa's founder, which will be discussed in more detail in the next chapter.

based on clearinghouses and correspondent relationships. In effect, the Fed offered a new type of clearinghouse, similar to those discussed earlier, but now with a national scope. It was also free of charge to Fed members, as the cost of clearing was covered by the Fed itself, which in turn was funded through taxes. Just as in the other clearinghouses, the Fed mandated that participating banks clear each other's cheques at par with the hope of completely eliminating non-par cheque clearing. However, participation in the cheque clearinghouse was voluntary at this point and only a quarter of the member banks chose to use the system. As Fernelius and Fettig observed "contrary to the Fed's hopes, most banks didn't care to join the Fed in its attempt to streamline the nation's payments system because there was too much to lose by such efficiencies."[23]

The reaction to the Fed's cheque clearing system provides a nice example of one of the core theories from the Social Shaping of Technology position: the choice to adopt a new, more efficient technological system is not always a given. From an engineering perspective, system efficiency is its own justification, and engineers are often perplexed when more efficient technologies are not universally adopted. However, when we take a historical perspective, we can see that increased efficiency is not always a benefit to all groups within the system. If increases in efficiency will harm a powerful group, that group will resist that change or even shape the system in such a way as to heighten the inefficiency. In the case of payment systems, float creates a powerful economic incentive for those that find themselves on the paying side of transactions more often than the receiving side. That group tends to resist changes to the payment system that would make it more efficient, reducing the time it takes to settle a transaction, thereby reducing the float. Conversely, those that would benefit from an increase in efficiency will fight for those changes, and the resulting power struggle will shape the development of the system.[24]

The Fed then changed the rules in April 1916, requiring that all member banks remit cheques presented by a Federal Reserve bank at par, regardless if they were presented in person or not. The Fed then exerted pressure on non-member banks to clear cheques coming from the Fed system at par. For those that refused, it presented their cheques in person to avoid discount fees. Litigation, legislation and industry changes caused the number of non-par banks to fluctuate over the next six decades, but by 1972 all states had finally outlawed the practice of discounting cheques, enforcing a par-clearing system nationwide.[25]

While there is some disagreement as to whether the Fed's cheque clearing system is indeed more efficient than the traditional correspondent and clearinghouse arrangements (which are still used today, albeit to a lesser degree), the elimination of non-par cheque clearing created the expectation that the processing of these kinds of payment instruments should be free, especially

[23] Quote from Fernelius and Fettig. For historical points, see Evans and Schmalensee, pp. 41–42.

[24] For another example based on the choice between water and hand milling, see Marc Bloch, 'The Watermill and Feudal Authority', in: Donald MacKenzie and Judy Wajcman, editors, *The Social Shaping of Technology*, 2nd edition (Maidenhead: Open University Press, 1999).

[25] Evans and Schmalensee, pp. 41–42; Fernelius and Fettig.

to merchants (i.e., the depositors). This created not only a preference on the part of merchants for payment by cheque as opposed to payment methods that incurred a discount (as credit cards would), but also a hostility to paying a fee to process anything that resembled or was marketed as a replacement for a cheque. We will return to this point in Chapter 9 when we discuss the creation of Visa's debit card product and the subsequent lawsuits against its pricing brought by merchants.

3.2.2 Merchant and Industry-Specific Charge Card Systems

Shortly after the Federal Reserve System was established, the payments industry experienced another set of innovations that can be grouped roughly under the heading merchant or industry-specific charge card systems. These systems were offered by specific large-scale merchants, or by a network of merchants within a specific industry or geographic area, and they allowed consumers to purchase goods and services on short-term credit, requiring full payment at the end of the billing cycle.

Customers generally did not pay a fee to have or use these cards, nor did they pay interest on the short-term credit extended to them, so the systems themselves tended to lack profitability. However, profit was not the purpose of these systems. Instead, merchants offered these systems for two other reasons. First, they believed that customers purchasing on credit tended to buy additional or higher-priced items than those who were paying with cash. Second, they believed that the cards tended to create a loyalty, or brand affinity, for those businesses that would otherwise be competing strictly on price. The cards not only carried a reminder of the merchant's name, they also lent a certain amount of prestige and recognition to the cardholder, which in turn encouraged the cardholder to shop at the merchant's store to the exclusion of others.[26]

What made these systems innovative was not the idea of buying items on credit. The idea of "buy now, pay later" was nothing new at the time. Mandell reminds us that "retailers drifted into the credit card industry simply by conducting business as usual. In the United States, a large proportion of retail sales, particularly of discretionary items, had always been made on credit."[27] What made these systems innovative was the ability to use a card, or similar device, to identify one's account to a centralized credit system that was available at multiple locations, or multiple merchants within an industry-specific network.

[26] Mandell, *Credit Card Industry*, p. 18. Note that the loyalty reason actually seems a bit tenuous. If a cardholder could obtain a card from one merchant, that same cardholder could likely obtain cards from competing merchants as well, as there was no centralized credit reporting agency that would stop a consumer from obtaining multiple cards. If cardholders tended to be more loyal to a given merchant, it may be for more complex reasons.

[27] Ibid., p. 17.

Western Union and the Department Stores

Western Union was one of those large-scale merchants that offered credit accounts to both businesses and recurring customers. Western Union also had a centralized accounting and billing office so that customers could charge their telegrams at any one of Western Union's numerous locations, but they needed a way to identify the customer to the billing system. In 1914, Western Union issued what is believed to be the first consumer charge card—a small rectangular piece of paper, containing the account number, the name and address of the person or company responsible for paying the charges, and a signature line.[28] The card identified the billing account, and the signature could be used to authenticate the cardholder.[29]

There were actually two types of cards, those for government officials and employees, and one for "representatives of business concerns and other responsible customers." Either type could be used to charge telegrams, as well as other services, at any Western Union location in the United States or Canada. The clerk would record the card number on the message, or a separate "toll ticket," and submit that to the central accounting office for billing.[30]

Department stores soon followed with cards or other tokens of their own.[31] Typically these cards were issued only to wealthier customers, who embraced the cards not only because they simplified their payments, but also because they lent a certain amount of prestige.[32] These cards were again only charge cards, so having one meant that you could pay off an entire month's purchases at once. The loyalty these cards created, or the merchant's perception of that loyalty and its value, became an important dynamic in the history of payment card systems.[33] As we shall see, the major department store retailers in America refused to accept bank-issued credit cards for many years, for fear of loosing this loyalty. Visa recognized this very real concern, and eventually developed the idea of co-branded loyalty cards that offered customers incentives to shop at specific merchants.

The early department store cards required clerks to hand-copy the account information from the card to the sales draft. This of course introduced the possibility of copying the number incorrectly or illegibly. To address these problems, the Farrington Manufacturing Company of Boston developed their "charga-plate" system, which was widely adopted by department stores starting in 1928. These were reconfigured metal address plates like the kinds used in

[28] Donald Jutilla, *Fundamentals of Bank Credit Cards* (Tacoma, WA: American Institute of Banking, Tacoma Chapter, 1973), pp. 8–10.

[29] It is perhaps not accidental that Western Union was at the forefront of payment systems innovation, as they also introduced the first electronic money transfer service in 1871.

[30] Western Union Telegraph Company tariff manual, provided through personal correspondence with Harold Smith, former operations employee and webmaster of Western Union Retirees web site, and Carlos van Orden, another former Western Union employee. Note that Chutkow incorrectly states that these cards were made from embossed metal, like a military dog tag. Chutkow likely confused this card with the "charga-plates" discussed later in this section. See Chutkow, p. 58.

[31] One early department store "card" was actually a small elliptical metal fob meant to be put on one's key chain.

[32] Mandell, *Credit Card Industry*, p. 18.

[33] Again, the loyalty argument may have had more to do with how merchants *perceived* the potential loss of loyalty rather than any real consumer affinity.

addressing machines of the time. They were small metal rectangles embossed with the account number, customer name, and address on one side. A bit of paper was affixed to the other side that identified the issuing store and contained a line for the customer's signature. These plates were then used with a new device called an *imprinter*, which would squeeze the plate against a new standardized sales draft. This draft contained an additional carbon paper layer, which automatically transferred the embossed account details on the charga-plate to the top layer of the draft.[34] This method was not only a more reliable way to transfer the customer's identity to the sales draft, but also an important precursor to making the drafts machine-readable in the future.[35]

Up until this point, retail charge cards were issued by each particular merchant, and could be used only at that merchant's locations. However, in the early 1930s, a new model began to emerge—several smaller retailers, who were normally competitors, banded together to form cooperative payment card systems. In these systems, a cardholder could use the same card (often a charga-plate) at multiple merchants. In some systems, the central organization would extend the credit and collect from the customer via one bill, while in others the central organization would simply issue the plates and the individual merchants would bill customers separately.[36] These networks of competing retailers were often quite extensive. For example, the Retail Service Bureau of Seattle included over a thousand retailers by 1936.[37] These organizations were similar in concept to Visa, though they were owned by the retailers and not the banks, and were limited to a specific geographic area.

The Oil Industry

In the early 1920s, charge cards began to spread to the oil industry. In 1924, General Petroleum Corporation of California began issuing what they called "courtesy cards," and other gasoline retailers quickly followed suit.[38] These were charge cards, made from paper, that were handed out by station managers to repeat customers. With this card, a customer could charge gasoline and related purchases at any of the stations in the same franchise, which could extend over a considerable geographic distance. These programs typically operated at a loss, but the oil companies persisted because their prime concern was not profit, but brand loyalty. Gasoline is a relatively undifferentiated product, and there is little that ties a customer to one gasoline

[34] Mandell, *Credit Card Industry*, p. 18.

[35] Mandell notes that when Standard Oil of California switched to metal embossed plates in 1952, errors in identification were cut by 94 percent and drafts returned to the stations due to the inability to establish identity were cut by 80 percent (Ibid., p. 23).

[36] Honey noted that centralized billing was not a widespread practice (Honey interview).

[37] Ibid., p. 18. Unfortunately, Mandell provides little detail about these cooperatives, and no other author mentions them. However, if they are the first instances of a set of competitors cooperating to provide a common payment system, they are of great historical importance.

[38] 'Facts About BankAmericard', National BankAmericard Incorporated Booklet (October 1975). See also Mandell, *Credit Card Industry*, p. 18.

station over another, especially while traveling.[39]

Any loyalty generated by oil cards helped to not only increase the franchise's sales, but also enroll new stations into the network. Oil companies could promise owners of new stations that they had a large number of loyal, card-carrying customers that would choose a station in their branded network over others. Due to the nature of automobile travel, motorists would often find themselves in unfamiliar places, and would naturally choose a gasoline station where they could use their card. Additionally, the card offered the new station manager a way to extend credit to his customers without the hassle of accounting, billing and collecting.[40]

However, the ability to leverage cards when signing new stations was dependent on a large cardholder base. In order to build that base quickly, these early oil cards were given out quite liberally. Mandell notes that the mere fact that a customer owned a car was often enough evidence for a station manager that the customer was credit-worthy.[41] In 1939, Standard Oil of Indiana went one step further, bypassing the station manager and distributing 250,000 unsolicited cards directly to consumers.

This episode provides us with the first illustration of the chicken-and-egg dilemma raised by multisided platform economic theory, which was mentioned in the previous chapter. Cardholders will not carry cards until there are enough merchants that will accept them, and merchants will not agree to accept cards until there are enough cardholders that want to use them. The oil companies' solution to this dilemma was simply to mass-issue unsolicited cards to consumers, and this pattern would be imitated by other payment card systems until it was finally outlawed in the United States in 1970. However, the benefits of this mass-issuance also came with a corresponding cost: large initial losses due to fraud and defaults. One man reportedly used his card to operate a long distance taxi service, charging over $500 worth of gas and oil without making any payments until he was finally arrested for an unrelated offense.[42]

The Airline Industry

The idea of the charge card then spread to the airline industry. In 1931, Century Airlines began issuing pre-paid coupon books for air travel, and after they were acquired by American Airlines, a Detroit office manager suggested to one of his large accounts that the airline could hold their coupons for them, and simply remove them from the books in response to "authorization letters" sent by the account. This program was widely adopted and eventually the books were replaced with a simple cash deposit. In 1936, the program was expanded to include other airlines and was renamed the Universal Air Travel Plan (UATP). UATP ran their own private clearinghouse to process the transactions, and by 1968 they were processing a billion dollars

[39] Mandell, *Credit Card Industry*, pp. 18–20.

[40] Ibid.

[41] Ibid., p. 19. Considering the price of gasoline at the time, the amount of credit exercised by any one customer would have been quite limited, so this rationale might not have been as unreasonable as it sounds.

[42] Ibid., pp. 19–20. Unless otherwise noted, all dollar figures are in US dollars.

in sales volume across 111,000 business accounts with 1.5 million cards in circulation.[43]

UATP is another example of a payment card system that included multiple *competing* organizations that cooperated just enough to provide a universal system.[44] The member airlines of UATP were fierce competitors, yet it was in all of their interests to make airline travel easier, especially for business people. Business trips often required traveling on multiple carriers, and thus it was attractive for businesses to have one account on which they could charge all their air travel, regardless of airline. Eventually UATP morphed into a more general-purpose travel and entertainment card, extending into other travel-related merchant categories, including trains, busses, hotels and restaurants.[45]

The Great Depression and the credit restrictions imposed as America entered World War II put a halt to nearly all these systems.[46] However, these early programs developed a number of ideas and practices that would influence the design of future payment card systems including BankAmericard and Visa: the identification card, embossed account information transferred by imprinters, cooperative joint ventures amongst competitors to provide a universal system, and mass unsolicited issuance to build the cardholder base quickly.

3.2.3 Travel and Entertainment Card Systems

Shortly after the conclusion of World War II, the American payments industry experienced another round of innovations. A new type of payment card system was developed starting in 1949, commonly referred to as *travel and entertainment (T&E) cards* because they initially focused on supporting business people who travelled and entertained clients. However, these systems quickly grew to include many types of merchants that were only loosely connected to the activities of "travel and entertainment."

These systems differed from the merchant and industry-specific payment card systems in two important ways. First, instead of receiving your card directly from the merchant or a merchant cooperative, you now received your card, and its associated credit, from a third-party organization, one that was not owned by the merchants themselves. That organization also signed up merchants from many different industries to accept the card, allowing the consumer to use one card at a wide variety of establishments. Thus these systems are often called *generalized* or *universal payment cards*. On a periodic basis, the third-party organization would then pay the merchants for any charges made on its cards, and collect from the cardholders. Thus, the merchant was no longer in the credit business, and all bookkeeping, billing and collecting was performed by the third-party organization.[47]

[43] Mandell, *Credit Card Industry*, pp. 20–21. 1968 statistics are from Jutilla, pp. 53–55.

[44] Although the oil cards allowed customers to charge purchases as multiple stations, those stations belonged to the same franchise and thus did not compete with one another to any great extent.

[45] Ibid.

[46] Frederick M Struble, 'Bank Credit Cards and Check Credit Plans in the Nation and the District', *Monthly Review of the Tenth Federal Reserve District* (July–August 1969); .

[47] Mandell, *Credit Card Industry*, p. 26.

Second, these systems were operated specifically for profit. As noted earlier, merchant-specific payment card programs were offered primarily to increase customer loyalty, and nearly all operated at a loss because the merchants did not charge any interest on the credit they extended. The T&E systems used a different business model. They "bought" the sales drafts from merchants at a discount, much like the discount on cheques discussed earlier, which was commonly six or seven percent of the total. For example, if a hotel charged a business traveler $100, the T&E organization would pay the hotel $93 but collect the full $100 from the cardholder. The $7 difference would cover their costs and hopefully return a profit. Initially the cards themselves were issued without charge to build a significant cardholder base, but the various T&E organizations quickly discovered that they also needed to charge cardholders a small yearly fee in order to remain profitable, especially as the cost of funds increased.[48]

Diners Club

In 1949, Frank MacNamara and Ralph Schneider formed the first T&E payment card system, known as Diners Club.[49] MacNamara's initial plan was to create a card that would be honored at restaurants all over New York City, hence the "Diners" part of the name, but it quickly expanded to also include everything from hotels and car rental agencies to florists and charm schools.[50] It became an extremely successful system, boasting 1.3 million US cardholders at its peak.[51]

The original card itself was a rather unusual design. It was made from cardstock, shaped roughly like a typical business card, with the account name, address, and number printed on the front, along with a signature line. However, the "card" was actually a small booklet that opened up to list all the establishments that accepted it. Initially, this was a relatively small list so it fit on just a few pages, and as the number grew, the list was separated from the card.[52]

The origin of the Diners Club idea requires some comment, as it has been inaccurately reported for many years and in many sources. The story reported by the press was that MacNamara was entertaining a business acquaintance one evening in a New York restaurant, and when the bill arrived, he discovered he had not brought enough cash. To avoid embarrassing himself in front of his guest, he telephoned his wife who then drove from Long Island to deliver more cash. While he waited, he realized that a responsible business man should have a card with which he could charge a restaurant bill anywhere in the city.[53]

This story would normally lend credibility to the theory that technologies develop due to

[48] Mandell, *Credit Card Industry*, p. 26.

[49] Some sources put the formation at 1950. It's likely that the company was formed late in 1949 and did not actually start operations until 1950.

[50] 'For Everything', *Time* (23 March 1959).

[51] Matty Simmons, *The Credit Card Catastrophe: The 20th Century Phenomenon That Changed the World* (New York: Barricade Books, 1995), p. 106.

[52] Ibid.

[53] Evans and Schmalensee repeat this story, but note that the original *Newsweek* article describes MacNamara as being at a coffee shop and forgetting his wallet altogether. Later reports tell the story as presented here.

pressing social needs, but unfortunately, it was completely fabricated. Matty Simmons, the first press agent for Diners Club, thought that the real story lacked sufficient glamour, and invented the near-embarrassing incident, no doubt partly to encourage business people to avoid the same mistake by applying for a Diners Club card. Simmons admits:

> My story has appeared literally thousands of times in newspapers and magazines and books all over the world. In a beautiful four-color brochure, the Diners Club still evokes the legend of MacNamara sitting in that restaurant—his dinner partner long gone—waiting for his wife and dreaming of a credit-card world.
>
> Of course, they couldn't know until reading this that the scene never took place.[54]

In reality, MacNamara simply thought of the idea one day and eventually decided it was worth trying.[55] Although it would be an overstatement to say that Diners Club was a solution looking for a problem, business people were not directly clamoring for a universal charge card. Simmons himself doubted that it would work.[56] The idea was not immediately obvious—it had to be explained and sold, both to consumers and merchants. Simmons described how MacNamara pitched the original idea to him:

> "It's perfect for the businessman," he enthused. "You pay all your business entertainment with one monthly check. You don't have to carry a lot of cash with you, and, most importantly, you have receipts for reimbursement from your firm and entertainment receipts for tax deductions."
>
> I listened unimpressed. He went on. "It's all of those things," he said, "but maybe most importantly, it's prestige. It's having a good enough credit rating to own one of these cards so that restaurants will treat you like somebody who's somebody. If you're a businessman and you're entertaining clients or working on a deal, the people with you will be impressed."[57]

This quote nicely summarizes how the T&E payment cards were marketed to business people, and many of these same points would be made by the bank-issued credit cards. Marketers stressed that the card was more convenient than cash, which could run out or be stolen, and more reliable than cheques, which might not be honored out of town. The card also allowed the business person to delay and consolidate their actual payments, which could easily mount when traveling. The card provided automatic bookkeeping for not only corporate reimbursement, but also income tax deductions.[58] But these business functions were minor compared to the *cultural meaning* of the card: *prestige*. At this time, cards were still issued rather selec-

[54] Simmons, p. 26.

[55] Ibid.

[56] In a strange twist, MacNamara himself left the organization after only three years because he concluded it was only a fad that would quickly fade.

[57] Ibid., pp. 21–22.

[58] This significance of this point is often missed from our current vantage point. Restaurants at this time did not provide receipts unless asked for, which was required in order to be reimbursed or to deduct business entertainment expenses on tax returns. The T&E cards provided that proof automatically.

tively, so not everyone could qualify for one. To have one meant that you, and consequently your organization, were worthy of what amounted to a blank cheque.

Diners Club built their cardholder base in much the same way as the oil industry. They initially sent unsolicited cards to "several thousand prominent businessmen with a letter revealing its wonders."[59] Simmons then began generating publicity, getting stories in every major New York newspaper. He also began advertising in newspapers and magazines, and eventually through direct mail. After the public was educated, applications began to flow in. In just one year, membership grew to over 100,000, and a similar number had applied but been rejected as bad credit risks. Initially, there was nothing to lose by applying because the card itself was free, but by late 1950, they were forced to charge a three dollar membership fee in order to generate additional revenue. Simmons writes that although they were concerned that this fee would destroy the cardholder base, it turned out to be a blessing, as the only customers who cancelled were those who were not actually using the card. Inactive users still create costs in a card system, but produce no revenue, so it is in the system's best interest to reduce their number.[60]

However, attracting cardholders was now only one side of the equation. The T&E card systems also needed to convince merchants to accept the card, which entailed giving up six or seven percent of the sale to the card organization. Diners Club sold the card to merchants using a variety of tactics that would be repeated throughout the history of bank-issued credit cards. First, they commissioned a study that claimed a Diners Club cardholder typically spent 18 percent more than a cash customer, which would more than make up for the discount.[61] Second, they used the merchant's competitive position either as a carrot or a stick: if the competition had not yet agreed to accept the card, they sold the merchant on the additional revenue they would get from the lucrative business traveller if they accepted the card; however, once one merchant agreed to accept the card, they could then pressure the competing merchants by warning them that they would lose sales to their competitor if they did not accept the card as well.[62] Diners Club had also started a magazine that was sent to all cardholders, and they often promised full page advertisements in order to close the deal with a merchant.[63]

However, some merchants, especially those that had their own industry-specific cards, actively resisted T&E cards. For example, the US Hotel Trade Association had their own program, called the Universal Travel Card, and actually prohibited their members from accepting

[59] Simmons, p. 27.
[60] Ibid., p. 29.
[61] Ibid., p. 41.
[62] Simmons describes such an episode when trying to break into the airline industry. After signing the newly-created Northeast Airlines, he called the president of Eastern Airlines. "'I just signed Northeast Airlines to a Diners Club contract,' I told him. 'Do you want all of our one million businessmen and travelers to use only Northeast on your common routes?' A week later, I signed Eastern. The rest of the airline industry fell into place soon after." (p. 48–49).
[63] Ibid.

Diners Club and other third-party cards that carried a discount.[64] In 1957, the major oil companies (with the exception of Texaco), banded together to block all third-party cards.[65] The airlines also initially refused to accept third-party cards like Diners Club, as their own card system required a healthy $425 deposit from each of their 800,000 members, which provided them with an enormous amount of interest-free working capital.[66] Large department stores also refused to accept the card—ironically, even though Alfred Bloomingdale had taken over as president of Diners Club in 1952, he could not even sign the large department store his family started.[67]

There were also active revolts by merchants who initially agreed to accept the card, but wanted the discount reduced or eliminated. For example, in the late 1950s, a group of restaurants in Milwaukee, and a similar group in Seattle, banded together and demanded that Diners Club reduce the discount fee or they would leave the system *en masse*. Their highly-publicized demands threatened to tip the balance of power in the Diners Club system, potentially destroying it. Simmons recounts how he dealt with the problem:

> I flew to Milwaukee with one plan in mind: divide and, hopefully, conquer. The director of the Milwaukee Restaurant Association invited me to speak at a specially called meeting. I arrived two days early and spent those two days meeting privately with owners of the leading restaurants. I offered advertising and publicity in the *Diners Club Magazine* and full-page ads in the Milwaukee papers.... In some cases, I agreed to advance needed funds against future charges by cardholders.[68]

When the vote was taken at the meeting, all but two of the restauranteurs agreed to stay in the system without a cut in rates. The news coverage of the vote spread throughout the country, and the Seattle restaurants quickly backed down.[69]

As Diners Club built its US merchant base, it also expanded internationally through franchise agreements.[70] It is not clear if UATP or Diners Club was the first international payment card, but Diners was clearly the first in the more generally-accepted T&E category.

[64] Mandell, *Credit Card Industry*, p. 27. Like the other industry-specific cards, The Universal Travel Card was not profitable and was eventually taken over by American Express when it issued its card in 1958.

[65] Ibid.

[66] Ibid., p. 28.

[67] Ibid., p. 7.

[68] Simmons, p. 41.

[69] Mandell provides a slightly different version of the Seattle story, based on his interviews with Alfred Bloomingdale. Bloomingdale claims that he broke the Seattle resistance by opening up his own restaurant that of course accepted Diners Club. Business people flocked to Bloomingdale's restaurant, forcing the others to break ranks and resume accepting the card at the original discount rate (Mandell, *Credit Card Industry*, p. 9).

[70] There is some debate over which was the first franchise. Mandell (relying on Bloomingdale) claimed it was the UK (p. 7), while Simmons recalls that it was France in 1955 (p. 37).

American Express

Diners Club enjoyed little competition throughout most of the 1950s and its yearly charge volume grew to nearly half a billion dollars.[71] However, in 1958 two important competitors entered the T&E card market. The first was American Express (AmEx), a company that had been in the payments industry since it invented the travellers cheque in 1891. Compared to charge cards, travellers cheques were a virtually risk-free business which provided AmEx with revenue not only from selling the checks, but also from the float realized on their cash reserves in between the time the travellers cheques were purchased and used. Many travellers cheques were (and still are) cashed long after they are purchased, and some never at all. Still AmEx saw the amazing growth and further potential of Diners Club and initially offered to purchase the company outright. According to Simmons, a deal was reached, but AmEx's chairman Ralph Reed thought the price was too high and decided AmEx should develop its own system.[72]

AmEx designed a system that was almost identical to Diners, but their card was made of a relatively new substance: plastic.[73] These new plastic cards were more durable than those made from cardstock, yet they were still lightweight. Most importantly, they could be embossed just as the old metal charga-plates were, allowing them to be used with imprinters for a more accurate transfer of the account information.

AmEx was an enormous company compared to Diners Club, and they proved to be a formidable competitor. In October, they mailed out eight million membership applications and ran full page advertisements in twenty-three newspapers across America. They used their relationship with the banks that sold their travellers cheques to place application forms in front of every customer. They also had an established network of travel agency offices around the world, from which they could sign up both cardholders and merchants. Their massive cash reserves allowed them to spend more on advertising in their first year than Diners Club had spent in their entire history.[74] However, the furor created by the AmEx advertising actually helped increase demand for Diners Club as well. By 1960, Diners Club membership was over the one million mark, but AmEx was close behind with 700,000.

However, as we shall see repeated throughout the history of payment cards, mad rushes to acquire cardholders are often accompanied by the lowering of credit standards, and consequently, a rise in credit losses. Diners Club had always enjoyed very low losses, partially due to their strict credit controls, but also because their cardholders were mostly business people backed by their corporations. During their membership battle with AmEx, Diners Club's credit losses rose from ¼ percent to almost one percent, and AmEx's losses were suspected to be much higher.[75]

[71] Simmons, p. 53.
[72] Ibid., pp. 61–66.
[73] Ibid., p. 82.
[74] Ibid., p. 70.
[75] Ibid., p. 75.

American Express took a large leap ahead when they computerized their bookkeeping and billing in the early 1960s. Up until this point, both American Express and Diners Club processed sales drafts and billed customer using entirely manual processes. AmEx's head of card operations came from the Air Force, where he had been in charge of data processing during World War II, and he realized that AmEx could be more efficient, and therefore more profitable, if it converted to computers. Diners Club eventually followed suit in 1967, but Simmons recalls that it was done "in a state of confusion and ineptness of classical proportions," and actually contributed to their first year of net loss since 1951.[76] The original founders of Diners Club had both died by this point, and by 1970 AmEx surpassed Diners as the premier T&E card system.[77]

Carte Blanche

The second competitor of Diners Club in the T&E card market was Carte Blanche, which originally started as a private card of the Hilton Hotel Corporation, but was eventually spun off into an independent organization in 1958. Carte Blanche has always had the minority of market share in the T&E card business, and thus is often ignored in the literature. However, they did make one important contribution that should be discussed here: the gold card. In October 1972, Carte Blanche began issuing a special gold-colored cards to their most credit-worthy members.[78] Mandell does not mention if this card offered any special privileges or benefits beyond expiring in two years instead of one, but it clearly created a new level of prestige. By the early 1970s, the number of cardholders had grown to such an extent that the card no longer carried the prestige that MacNamara thought would be its best selling point. The creation of a new tier of card allowed Carte Blanche to grow their cardholder base while still offering prestige to their most profitable customers.[79]

3.2.4 Bank Credit Cards

Shortly after the launch of the American Express card in 1958, the American payments industry saw another series of innovations from which we can more directly trace the origins of Visa. These new payment card systems were created and operated by an actor group that had been conspicuously absent throughout the early years of payment cards: the banks.

Beyond having a new type of issuer, these new systems differed from the T&E systems in four important ways. First, they were focused more on middle-class consumers than on business people, and thus enlisted a different and more diverse set of merchants. Second, they

[76] Simmons, p. 103.
[77] Ibid., p. 106.
[78] Mandell, *Credit Card Industry*, p. 110.
[79] Mandell may be incorrect here. He goes on to claim that American Express first offered their gold card in 1975 (p. 111), but other sources such as the AmEx web site say they issued their gold card starting in 1966. There is also no mention of a Carte Blanche gold card in the 1972 index of the *American Banker*. Therefore, it may have actually been American Express that issued the first gold card and not Carte Blanche.

offered a feature that some merchant-specific systems had recently developed, which would appeal to those consumers: revolving credit.[80] Cardholders could choose to pay the entire bill at once, or finance any portion of their purchases, repaying them over several months with interest. Third, as opposed to the T&E issuers, the banks offered their cards without any annual fees.[81] Fourth, merchants received payment for their sales drafts, less the normal discount, immediately upon deposit, as opposed to several days later as was common in the T&E systems.

Early Bank Charge Card Systems

Although the systems discussed in this section began in 1958, it should be noted that some smaller banks had actually experimented with the charge card model (i.e., no revolving credit) even before the creation of Diners Club. Flatbush National Bank of Brooklyn, New York issued their "Charg-It" card in 1947, Patterson Savings and Trust of New Jersey issued a similar card in 1950, and Franklin National Bank of Franklin Square, New York began their system in 1951.[82] The idea quickly spread, and during the next four years, over 100 banks started similar charge card systems. However, at the end of that period, only 27 of those programs were still operating, and no new programs were started for the next two years.[83]

There are two main reasons why these early programs failed. First, they were unprofitable because the banks did not charge the cardholders an annual fee, as Diners Club did, and they lacked the possibility of revenue from interest charges on a revolving credit line. The reasons for not charging a cardholder fee are not discussed in the literature, but one can surmise that the main competitors for these systems were the merchant and industry-specific card systems mentioned earlier, which did not charge a fee. These competitors were also unprofitable, but as explained earlier, loyalty rather than profit was their prime objective.

These early experiments also failed because they could not build the critical mass of adopters necessary to sustain a payment card system. These types of systems have large startup and operating fixed costs that can only be offset by a large volume that brings about an economy of scale. Additionally, multisided platforms tend to have a tipping point with respect to participation—if there are not enough participants on either side, people start to lose interest and the system

[80] According to Mandell, The L. Bamberger and Company department store was the first to develop a system with revolving credit in 1947 and it was quickly adopted by the major New York stores such as Gimbels and Bloomingdales. In 1956, J. L. Hudson's of Detroit added the idea of an interest-free period. However, not all merchant-specific systems offered revolving credit, and the T&E systems were prohibited by law from doing so, so the bankcard systems starting with BankAmericard are often categorized as a separate set of innovations because revolving credit was a common feature. See Mandell, *Credit Card Industry*, pp. 24–25.

[81] Struble, p. 5.

[82] Mandell, *Credit Card Industry*, p. 26; Struble, p. 4. Note that many sources claim that Franklin National Bank was the first "credit card," but a close reading of Struble reveals that although he used the that term, the Franklin plan did not offer revolving credit and was thus similar to the T&E cards. Struble wrote in 1969, and the terms defined above were not yet used consistently. Other sources have suggested that Franklin first offered a non-revolving card, and then later added the revolving credit feature, thereby becoming the first bank credit card.

[83] Ibid., p. 5.

collapses.[84] Due to the way banks were regulated in the United States at this time, most banks could not build this critical mass. Banks were prohibited from operating across state lines, and most states did not allow banks to open more than a few branches. Some states, called *unit banking* states, restricted banks to just one physical location. This severely limited the number of customers to which a bank could issue cards, as well as the number of merchants they could enlist to accept the cards. In fact, the "Charg-It" card was accepted only in a two-block radius around the bank itself.[85] For the car-loving consumers of the early 1950s, a two-block radius was simply not enough to sustain interest.

The BankAmericard

However, there was one bank in the country that had the resources, potential scale, and the corporate culture, to make it work: The Bank of America (BofA). BofA operated in California, a state which allowed banks to operate branches statewide.[86] In the late 1950s, California was also one of the most populous and wealthiest of states, and BofA had a banking relationship with 60 percent of its residents and held more than 30 percent of its deposits. With assets of $5 billion, BofA was not only the nation's largest bank, but also one of the largest in the world.[87]

As opposed to other large banks, BofA was also *culturally* predisposed to develop a credit card for the middle-class consumer. Most large banks of the 1950s did not engage in the extension of consumer credit, much less unsecured consumer credit, which was considered to be the domain of merchants and the less-than-reputable finance companies. Mandell writes that "If a bank had a consumer loan department, it was often found in the basement where no one could see the furtive borrower."[88] One critic of these early bank credit card systems complained that they were "lowering banking's image by engaging in an activity more properly associated with pawn shops."[89] Most bankers preferred to deal with safer, larger and more lucrative commercial loans. BofA, however, had a different organizational culture. Nocera explains that "It was a bank with the mentality of a finance company and proud of it."[90] It was started by A. P. Giannini, the son of an Italian immigrant, who prided himself on serving "the little fellow."[91] Ken Larkin, the BofA executive who would become synonymous with the BankAmericard program, saw it as a natural extension of the bank's business: "We were always a leader in installment credit. Anything you could buy on time we financed...the credit

[84] However, Evans and Schmalensee remind us that this tipping point does not necessarily mean that once the point is crossed, participation is entirely self-sustaining. A small change in pricing can easily cause consumers to leave a system, resulting in a quick collapse (e.g. Yahoo bids). See Evans and Schmalensee, pp. 134–139.

[85] Mandell, *Credit Card Industry*, p. 26.

[86] Nineteen states and the District of Columbia allowed statewide branching, sixteen allowed limited branching, and fifteen enforced unit banking. See Lawrence G Goldberg, 'The Effect of State Banking Regulations on Bank Credit Card Use', *Journal of Money, Credit and Banking* 7:1 (February 1975).

[87] Nocera, pp. 18–20.

[88] Mandell, *Credit Card Industry*, p. 29.

[89] Quoted in Nocera, p. 24.

[90] Ibid., p. 17.

[91] The history and culture of BofA is chronicled in a number of sources. See especially Chutkow and Nocera.

card was just a natural extension of that."[92]

However, The BankAmericard, as their system was called, was a slight departure from the traditional consumer installment loan. Traditional loans were secured, meaning that the item being financed could be repossessed by the bank if the consumer defaulted on the loan. The BankAmericard credit line was unsecured, and the bank had little recourse if a consumer could not pay. Traditional loans also required the consumer to apply for the loan, which would entail a review of the consumer's credit position, and their intended use of the funds, by a bank loan officer. With the BankAmericard, cardholders could finance anything they wished without ever visiting the bank. It was a form of self-service credit that transferred the financing decision from the bank to the cardholder. In effect, it subtly blurred the line between buying and borrowing.[93]

The designer of the BankAmericard system, Joseph Williams, had friends at Sears and Mobil Oil, and he patterned his new system directly upon theirs.[94] The card was offered to consumers without charge. Upon receiving their bill, cardholders could pay the entire amount and not incur interest, or pay less than the total and finance the rest at the rate of 18 percent per year. Merchants were charged a six percent discount fee on their transactions, but would receive funds immediately upon deposit of the sales drafts without the need to bill or collect from the cardholders. Merchants also paid $25 a month to rent a card imprinter.

The card itself was made of plastic with embossed account information, similar to the new American Express card. Just as in the case of the charga-plate, the embossed information was transferred to the sales draft using an imprinter, which reduced the potential of copy errors. Early imprinters did not have wheels for transferring the transaction date or amount, which had implications on machine processing that will be discussed in later chapters, so merchants hand-wrote these details on the sales slip, and customers added their signature to authorize the charge.[95]

The BankAmericard also had a formalized authorization process that was based upon the department store systems, but expanded to a multi-merchant environment. Each merchant was assigned a *floor limit*, over which the merchant was required to call for authorization.[96] This authorization process will be discussed in more detail in the following chapter, but it should be noted here that at this time it was entirely manual and exceedingly slow.

The accounting side however was computerized from the very beginning, albeit in rather limited way. BofA was actually the first bank in America to use a computer, an IBM 702 installed in 1955, upon which SRI developed a program to automate BofA's demand deposit

[92] Chutkow, p. 61.

[93] Nocera.

[94] Ibid., p. 24.

[95] Merchants were also supposed to fill in details about the goods purchased, but most found this too time consuming and neglected to do so.

[96] The term *floor limit* comes from the department stores, where it literally meant the amount under which the "floor" could authorize. Any amount above require a telephone call to the finance department (Powar interview).

accounts.[97] To support the BankAmericard, BofA adapted this computer system to maintain cardholder accounts and process sales drafts. Instead of using magnetic ink as they did with cheques, the sales drafts themselves had a punch-card as the bottom layer, which would be punched with the transaction details upon deposit.[98]

"The Drop"

Nocera writes that BofA approached the rollout of their system cautiously. They chose to test it in the relatively isolated town of Fresno, California, partly because they had a banking relationship with 45 percent of the families there, but also because they were not entirely sure it would work. They reasoned that if the card failed there, the bank's reputation would be damaged less than if the test was conducted in San Francisco.[99]

Just as in the case of Diners Club, the BankAmericard system faced the classic chicken-and-egg dilemma. Convincing merchants to give up six percent of the transaction and rent an imprinter for $25 a month would be possible only if there were a significant number of cardholders wanting to use the card. BofA solved the dilemma in much the same way all the previous systems did. They referred to it as "The Drop." In September of 1958, they simply mailed 60,000 unsolicited cards to households in Fresno. They followed this up with aggressive advertising to educate those consumers about their new card.[100]

As in the other systems, merchants faced a central tension when deciding whether to participate. Accepting the card meant possibly losing the customer loyalty built by the merchant's private card system, but not accepting the card put the merchant at risk of losing customers to a competitor that did. In general, the larger merchants with established card systems, such as Sears, JC Penney and Wards, did not accept the BankAmericard, but the smaller merchants did. These smaller merchants considered the loyalty to be less valuable than the costs of maintaining, billing and collecting on credit accounts. Ken Larkin recalled one merchant he visited:

> He had three girls working on Burroughs bookkeeping machines, each handling 1,000 to 1,500 accounts. I looked at the size of the accounts: $4.58. $12.82. And he was sending out monthly bills on these accounts. Then the customers paid him maybe three or four months later. Think of what this man was spending on postage, labor, envelopes, stationery! His accounts receivables were dragging him under.[101]

[97] The system was called Electronic Recording Machine-Accounting (ERMA). That system automated the processing of cheques through the use of magnetic ink printed on the back of the cheques, a technique that would eventually be used in the Magnetic Ink Character Recognition (MICR) standard in 1957. According to O'Brien, the first ERMA installation was demonstrated in San Jose in early 1956, though the BofA did not fully convert all their accounts to ERMA until 1962. See James A O'Brien, *The Impact of Computers on Banking* (Boston: Bankers, 1968), pp. 2–5; and Martin Campbell-Kelly, *From Airline Reservations to Sonic the Hedgehog : A History of the Software Industry* (Cambridge, MA: MIT Press, 2003), p. 49.

[98] Jutilla, Derman, and Russell interviews.

[99] Nocera, p. 25.

[100] Ibid., p. 15.

[101] Quoted in Ibid., p. 27. Stallwitz asked merchants in 1968 why they chose to accept the card. The reasons given

However, BofA needed to solve one more problem to make the system truly work: how would cardholders recognize merchants that accepted the card? Diners Club had simply provided a list of establishments to cardholders, but the number of merchants accepting the BankAmericard would potentially be far greater and more diverse. Cardholders needed a simple way to identify merchants that accepted the card without consulting a pre-printed list. Their solution was to create a *mark* for the system, which would be printed on both the cards as well as signs that hung in the windows of participating merchants. The mark looked a bit like a flag. It consisted of three colored bands—blue, white, and gold—running across the background, with the word BankAmericard written in white within the blue band.[102] This was a new twist on the function of identity discussed earlier. Not only would the card identify the cardholder's account to the system, the mark on the card would also identify a participating merchant to the cardholder.[103]

The Fresno drop went smoothly, though it attracted little attention, and consumers began to use the cards. Today it seems natural that consumers would want to use credit cards to pay for goods and services, but one should question *why* the people of 1958 Fresno would bother. Although the Diners Club card provided an obvious benefit to the traveling business person, the general consumer shopping locally would actually have little need for such a device. Cheques were commonly accepted at local shops, most merchants already extended credit to their frequent customers, and anyone who could qualify for a BankAmericard could no doubt qualify for a traditional installment loan. Considering the manual authorization process of the time, using the BankAmericard for purchases over the floor limit would have been more time consuming, awkward and embarrassing for the cardholder than writing a cheque. Nocera offers three possible reasons why the card was adopted. First the BankAmericard fit with the general trend in America towards impersonal self-service. Applying for a traditional installment loan meant looking a loan officer in the eye and promising to repay the loan. The BankAmericard arrived unsolicited and the decision to finance a purchase was entirely left up to the consumer. Second, the card did offer a level of convenience through consolidation. Instead of maintaining accounts and carrying cards from multiple merchants, the consumer could carry one BankAmericard and see all their charges on one bill. But his third reason is perhaps the most convincing—consumers used the card because it was a novelty in an age and culture of novelty. Nocera writes:

were similar to how the cards were marketed to merchants: it would increase their sales; reduce their risks and costs of extending credit; for competitive reasons; and for customer convenience. However, he noticed that residential merchants also tended to accept the card as a personal favor to their bankers, primarily because those bankers had given them their initial business loan. Stallwitz continued: "The most unusual reason for joining was put forth by one merchant who had just moved into a new store (not a related type of business) and a Master Charge sticker was already on the door. Rather than scrape it off, he signed up on his next trip to his primary bank." See Stallwitz, p. 44.

[102] Note that the mark started to appear sometime shortly after 1958. Images of the very first BankAmericard show that it did not have the blue, white, and gold bands. See Chutkow.

[103] Honey interview.

> The card was a novelty at first; just as Americans spent hours staring at the test pattern of their new TV, so did the citizens of Fresno gather around the checkout counter to watch someone pay with a BankAmericard. This was the 1950s, after all, a time of wonder at the miraculous march of progress. BankAmericard was part of that march.[104]

Shortly after the Fresno drop, BofA learned that their competitors were planning to launch a card system of their own in San Francisco.[105] Nocera writes that upon leaning this, all caution was put aside. BofA began sending cards to nearly all their depositors and consumer loan customers in California, and using their extensive network of branches to enlist as many merchants as possible.[106] Over the next 13 months, BofA issued 2 million BankAmericards and signed up over 20,000 merchants.[107] In one move, BofA created more cardholders just in California than Diners Club had ever had nationwide.

However, as the cards reached the urban center of Los Angeles, BofA began to experience the inevitable effects of sending out millions on unsolicited, unsecured credit cards. The problems of fraud in the BankAmericard system will be discussed in more detail in the next chapter, but the initial effects were staggering. Delinquencies were 22 percent, compared to 4 percent on traditional installment loans. Both thieves and merchants were creating numerous fraudulent transactions under the floor limits to avoid detection. Within 15 months of the Fresno drop, the BankAmericard system had officially lost $8.8 million, but Nocera estimates that the real losses were actually closer to $20 million.[108]

Despite these early losses, BofA chose to continue the system, reducing the number of outstanding cards and weeding out dishonest merchants. By May 1961, the BankAmericard system was generating a profit, though this was initially kept quiet as the publicity from the early losses was still helping to keep other banks from developing competing systems of their own.[109]

Other Bank Credit Card Systems

Although the BankAmericard was the most successful of the early bank credit card systems, it was not alone. The other system most often discussed in the literature came from Chase Manhattan, which started the Chase Manhattan Charge Plan (CMCP) in 1958. However, it

[104] Nocera, p. 28.

[105] According to *Business Week*, this was the First Western Bank and Trust Company of Los Angeles. See 'The Charge-It Plan That Really Took Off', *Business Week* (27 February 1965), p. 58.

[106] Larkin later remarked that this was "a calculated risk, done just the one time to get the plan off the ground." See Ibid.. Honey noted that BofA also used lists of charga-plate holders provided by the retail service bureaus (Honey interview).

[107] Nocera, p. 30.

[108] Ibid., p. 31.

[109] 'Charge-It Plan', p. 58. For the accusation that it was purposely kept quiet, see Galanoy. Mandell actually questions this profitability, claiming that BofA did not include the cost of funds or advertising in the accounting of the card system, thus making it look profitable when it actually was not (Mandell, *Credit Card Industry*, p. 30).

was never profitable and they were eventually forced to sell it in January 1962, partially because the large New York department stores refused to accept it.[110] Ironically, it was sold to Joseph Williams, the designer of the BankAmericard system, who had resigned after its initial losses.[111] The system was renamed Uni-Card, it was soon acquired by American Express, and finally repurchased by Chase in 1969. Mandell writes that "The second time around proved no more profitable than the first," and Chase shut down the Uni-Card and eventually became a member of the organization that would become Visa.[112]

In addition to Chase and BofA, 29 other banks started credit card systems of their own during 1958 and 1959. However, nearly all of these programs failed or reported massive losses during their initial years. The terrible press from these programs further discouraged other banks from starting systems of their own, and from 1960 to 1966, only 10 more banks created new systems.[113]

However, this trend then reversed quite suddenly. Between 1966 and 1968, 440 new systems were created, more than had ever existed before. Struble offers four possible reasons for this sudden explosion of new systems. First, by 1966 the rest of the industry finally became aware that the BankAmericard was not only making profits, but also reporting substantial increases in those profits each year. Profits were increasing primarily because the large number of cardholders were financing more of their purchases, which in turn generated substantial interest income. This good news emboldened some banks to start their own systems.[114]

Second, Struble claims that the availability of new "high-speed computers" significantly reduced the costs of operating a new card system. Indeed, the "third generation" IBM 360 line had recently been developed, and large banks such as BofA began installing them in 1966.[115] However, these machines were still prohibitively expensive for most medium or small sized banks, so although they might have contributed to the profitability of existing systems such as BankAmericard and therefore been an indirect cause, it is unlikely that they were a direct impetus for the creation of new systems.

Third, bank credit card systems afforded banks a new way to compete with one another in a time of strict regulation. In the 1960s, American banking regulations dictated the amount of interest a bank could offer consumers on savings accounts, and completely prohibited the payment of interest on demand deposit accounts. Banks had no mechanism with which to compete, forcing many to offer gifts such as toasters to those who opened new accounts. Credit

[110] Mandell, *Credit Card Industry*, p. 30; Jutilla, p. 44.

[111] 'Charge-It Plan', p. 58.

[112] Mandell, *Credit Card Industry*, p. 40. Mandell actually claims that Visa "acquired Uni-Card," implying that Visa purchased the portfolio, but this would be out of character for that organization, which was an association of card issuers and acquirers. Visa's chief of operations at the time confirmed that NBI did not purchase anything from Chase, and that Chase merely joined Visa and offered a Visa product (Russell interview).

[113] Struble, p. 5.

[114] Ibid.

[115] O'Brien remarks that the 360s were so much more powerful that they easily justified the cost of porting applications. In 1966 BofA replaced their 32 GE mainframes used for the ERMA system with just 2 IBM 360/65s (O'Brien, p. 14).

cards were a relatively new phenomenon and thus were not covered by the regulations. Banks used the credit cards to lure customers, either by offering a card when their competitor did not, or offering it with a lower interest rate. Once a bank did offer a credit card, all its competitors would no doubt respond by offering cards of their own.

Finally, Struble surmises that many banks created new systems around this time because they thought credit cards were the first step on the path towards the "cashless-checkless society." The wide-spread adoption of computers into banking in the 1960s brought about the idea that all value transfer could be accomplished electronically, substantially reducing operating costs for the banks. Although credit cards used paper sales drafts at this time, the back-office processing of credit card transactions was typically done using computers, and the idea of a merchant terminal that would eliminate the paper altogether was already being discussed in the late 1960s. If a banker thought that the "cashless-checkless society" might come to fruition, getting into the credit card business was the first step towards remaining relevant.

However, Struble's final reason leaves some room for criticism. At this time, credit card departments were the poor step-child of the bank, often located in the basement or other suitably out of the way location. If bankers seriously considered credit cards to be an important step toward a desired goal, they would have made the credit card department a more integral part of the bank's operations. It is also not clear that the majority of bankers thought an electronic value exchange system was such a good idea. Much of Visa's early efforts were directed at getting their member banks to support such an idea, and speeches by Visa's founder reveal a certain amount of frustration at the lack of support from his membership.

Surprisingly, Struble omits one other cause that could easily account for the sudden growth in new card payment systems: the formation of the two major national bankcard associations and their attempts to expand their network by enlisting new banks. We now turn our attention to these final innovations.

3.2.5 National Bankcard Associations

The final set of innovations to discuss in this pre-history of Visa are the two major national bankcard associations: The BankAmericard Service Corporation which would eventually transmute into the organization now known as Visa; and Interbank, which would eventually became the organization now known as MasterCard.

As noted earlier, by the mid 1960s the BankAmericard system had overcome its initial difficulties and was generating increasing profits, but it was still restricted to the state of California. BofA realized that both consumers and commerce were increasingly traveling across state lines, and for their card to be truly useful, it had to be accepted nationwide. American banking regulations at the time prohibited BofA from opening branches in other states, so they decided the best way to expand the system was to license the program to banks in other states.[116] BofA

[116]Note that BofA could have legally solicited cardholders outside of California, but it would have made little

created a subsidiary organization known as BankAmericard Service Corporation (BASC) that was tasked with signing up licensees and administering the entire system.

Licensee banks paid BofA $25,000 for the franchise, plus a percentage of their transaction revenues as a royalty. In return, they received the accounting software developed for the BankAmericard system, as well as an invitation to a training session in San Francisco.[117] However, much to the dismay of the licensee banks, this training session was given by the marketing department, and much of the discussion revolved around the marketing aspects of the program. Many of the licensees discovered that they could obtain more helpful and accurate information on how to run their programs by visiting the BankAmericard processing centers, directly observing and talking with their operations people.[118]

Initially BofA licensed only one bank in any particular geographic area, essentially providing it with a local monopoly.[119] These licensee banks typically had correspondent relationships with BofA, and thus were "loyal" or at least tied to BofA in some sense. Although this practice might seem a bit exclusive or restrictive, it was likely a necessary consolation in order to entice banks not only to pay the license fee and royalties, but also to give up their chance to issue a card with their own brand.[120]

The licensing system also created a new function never before seen in payment card systems: *interchange*. Because cardholders from one bank could now use their card to make purchases at merchants represented by a different bank, the two banks needed a way to clear and settle those transactions. In all previous payment card systems, the same organization both issued the cards and represented the merchants, so all settlement and clearing was done within the same organization. Cheque payment systems had always experienced this scenario, and as discussed earlier, the Federal Reserve had established a national clearinghouse in 1915 for just this purpose. However, in a move that may have perhaps sown the seeds of its own destruction, BASC chose not to create a centralized clearinghouse for the BankAmericard system. Instead, acquiring banks were required to mail their drafts directly to the issuing bank for payment, less a discount fee, now called the *interchange reimbursement fee*. This will be discussed in more detail in the next chapter, as it became one of the central reasons for the creation of a new independent organization.

The banks that competed with the BankAmericard licensee banks quickly reacted by forming regional, non-profit cooperative associations of their own. In many cases, these regional as-

sense to do so. There were no centralized credit reporting agencies at this time, so BofA had no way to establish the credit worthiness of prospective cardholders in other states without the help of a local bank. Furthermore, directly signing up merchants would have been extremely difficult, as the merchant would have had to maintain an account with a BofA branch in California.

[117] Nocera, p. 56.

[118] This was mentioned both in Hock's autobiography, as well as interviews with Don Jutilla, director of one of the first BankAmericard licensee programs. Jutilla created flowcharts from his observations that were then used to train his staff.

[119] These areas were also international—Barclay's Bank became the first international licensee in 1966, and the sole BankAmericard issuer in the UK.

[120] Initially banks were not allowed to add their brands to the card. See the next chapter for more details.

sociations were also centralized processors. The members still issued cards, signed merchants and held the receivables, but the regional association provided the more mundane operational functions such as authorization, sales draft processing, accounting, and billing. By centralizing their operations, these associations could also achieve an economy of scale, which reduced the operating costs for each of the members.

These regional associations then joined together into a national, non-profit cooperative association known as Interbank in order to allow their cards to be used across the country. However, national merchant acceptance was hampered by their lack of a common name on the card. Although all Interbank cards contained a common mark, it was only a small "i" in one corner of the card, barely noticeable compared to the regional association's name and marks, which varied from region to region. In contrast to the BankAmericard system, which used and promoted a consistent name and mark, the Interbank system did not actively promote their common mark, and thus cards were not as readily accepted outside their issuing region. In 1969, Interbank began to address this problem by purchasing the rights to the name first developed by First National Bank of Louisville, Kentucky, and the mark popularized by the Western States Bankcard Association (WSBA).[121] The name was "Master Charge" and the mark was the overlapping yellow and orange balls, and these were eventually used on all cards issued by Interbank members. In the 1980s, they changed names again to MasterCard and just recently they became an independent, for-profit stock corporation.[122]

3.3 Conclusion

In this chapter we reviewed the major innovations in the American payments industry, from the formation of the Federal Reserve to the National Bankcard Associations, in order to construct a historical context for Visa's formation and development. This context helps us understand when and where many of the ideas and techniques found within the Visa system originated: the identity function of the card; embossed cards and imprinters to reduce copy errors; mass-issuance of cards to prime the system; the techniques for educating consumers and enlisting merchants; revolving credit; gold cards; common marks to help cardholders identify participating merchants; and the introduction of computers for accounting and transaction processing.

This context also establishes many of the economic, social, and regulatory dynamics that shaped (and continue to shape) the history of Visa and its system: the tradition of discounts in the clearing process; the elimination of those discounts in the Federal Reserve's cheque

[121] Some sources mistakenly report that the name was created by WSBA. However, WSBA merely licensed the name from First National of Louisville, but WSBA did develop the mark, and made both widely recognized in the West (Honey interview).

[122] The story of Interbank and MasterCard is also beyond the scope of this dissertation, except where it intersects Visa's story. Strangely, there seems to be even less written specifically about Interbank and MasterCard than there is about Visa. Some history appears in Jeffrey Kutler, 'Metal Plates to Duality: The Shaping of an Industry', *American Banker* (9 September 1994); as well as Evans and Schmalensee; and Mandell, *Credit Card Industry*.

clearing system and the expectation of free processing for anything like a cheque; the loyalties (real or perceived) created by merchant and industry-specific card programs that are threatened by third-party universal cards; and the American banking regulations that required cooperative joint ventures for a nationwide system.

With this context established, we now turn to the formation of the Visa organization, and its first steps.

Part II

History of the VISA Payment System, 1970–1984

Chapter 4

Creating a System: Dee Hock and the Foundation of the Organization

Part I set the context for understanding the history of the Visa payment system that I will now present here in Part II. In this chapter, I will focus on the formation of the organization and its first steps, covering the period of 1968 to 1971.

By 1968, the BankAmericard licensing program appeared on the surface as if it was a resounding success. The number of licensee banks had grown from the initial 8 to 254, and each bank had further sub-licensed other agent banks in their territories. There were now 6 million cardholders, each of which could use their card at 155,000 merchants in 17 different states, as well as a few foreign countries. The system was processing $458.9 million in sales volume, and growing each year.[1] And then in October of 1968, it very nearly disintegrated.

4.1 Problems in the Licensing Program

What was actually wrong with the BankAmericard licensing program? Why did those involved consider the creation of an independent organization to be the only possible solution? This section will discuss the various problems faced by the licensing program and show that they were not likely to be fixed under the existing organization, primarily because the problems were both operational and organizational.

4.1.1 A Typical Transaction in 1968

To understand the various operational and organizational problems, we must first understand what it was like to initiate, authorize, clear, and settle transactions in the BankAmericard system of 1968. This is best done by walking through the process of a typical domestic purchase. Note that in this example, we will discuss only the details that will help us understand the specific

[1] 'Card Plans Show Big Gains', *Burroughs Clearing House* 52:5 (February 1968).

operational and organizational problems faced by the BankAmericard system; other interesting but less relevant details will be examined in later chapters. This example is also the ideal case; the unfortunate realities of the process will be noted in the next section.[2]

Imagine yourself in 1968, holding a shiny new BankAmericard. As discussed in the previous chapter, all cards had the same blue, white, and gold bands across the face of the card so that merchants could easily identify your card as acceptable, regardless of which bank actually issued it. Merchants also hung signs with the same marks in their windows, so that you could easily identify those that accepted the card, regardless of which bank represented the merchant in the system.[3] You spy a merchant that you need to visit, select your items and present your BankAmericard for payment.

If your purchase amount is below the merchant's floor limit the merchant can complete the transaction immediately without authorization.[4] The floor limit varied by merchant type and by card type: some cards had a star on the front, while others did not.[5] The floor limit for a general merchant was typically $50 for a non-starred and $100 for a starred card. However, airlines, hotels and other services were granted higher floor limits.

If your purchase is above the merchant's floor limit, the merchant is required to call for authorization. The merchant dials the acquirer's authorization center and verbally conveys the transaction details to the authorization operator. The authorizer first determines if the card was issued by the same bank or another by looking at the first four digits of your account number. If it is the same bank, the transaction is known as *local* or *on-us*, otherwise it is known as an *interchange transaction*.

If this is an on-us transaction, the authorizer then consults a series of printed reports to determine if the transaction should be authorized. At this time, there were no interactive computer systems with CRT terminals installed at the BankAmericard authorization centers. When your bank became a BankAmericard licensee, it did receive some "computer software" from the Bank of America, but this was just a simple punch card-based accounting system. This system produced two reports to help the authorizers: a list of known *hot cards*, which were either stolen or on hold for some other reason; and a summary of each cardholder's account, listing their current balance, credit limit, purchase and payment history. The authorizer first searches through the hot card list to ensure your account number does not appear there. Then the authorizer manually wades through the massive binder of account sheets to find yours, reviews your details, and consults the hand-written list of authorizations already given since the report was last printed. If all is in order, the authorizer gives the merchant an authorization

[2]Unless otherwise noted, the information in this section comes primarily from interviews with Jutilla, Russell, Honey, and Derman, as well as Jutilla; Fisher et al.; and Hock, *One From Many*.

[3]The merchant signs differed from the card in one strange way—the BankAmericard name was in the white band instead of the blue. This would be rectified later when the name of the system changed to Visa and the cards were redesigned.

[4]For a definition of the term *floor limit*, see page 44 in the previous chapter.

[5]This star created a kind of early segmentation. Jutilla remarked that his friends quickly noticed the difference and would often ask him how they could get a star on their card.

code, consisting of a few letters and digits, and the merchant writes that on to the sales draft.

However, if this is an interchange case, the merchant's authorizer does not have access to your records and is thus required to call or telex your bank's authorization center. The authorizer puts the merchant on hold, dials your bank's center and relays the transaction details. Your bank's authorizer then consults the same type of reports already discussed, and supplies an authorization code. The original authorizer then relays this code to the merchant.[6]

After authorization, the merchant then completes the sales draft. The draft is a multi-layer document: the top two layers are like tissue-paper, one for you and one for the merchant. The bottom layer is an IBM 80-column punch card, complete with the corner notch. The merchant puts your card and the sales draft into an imprinter, informally known as a "zip-zap machine," which squeezes the embossed characters on your card against the sales draft, thereby transferring your card number, expiration date and name onto each layer via carbon paper. The imprinter also holds another embossed plate containing the merchant's details. The merchant manually adds the transaction date and purchase amount to the draft, and you sign it to complete the purchase. The merchant is then required to check the signature on the card against your signature on the draft to ensure that you are the proper cardholder, but few do. The merchant tears off the customer copy and hands it to you, putting the other two layers in the cash register.

On a regular basis, the merchant deposits the punch card layer just like a cheque. However, unlike a cheque, the merchant receives an instant credit, less the discount, the amount of which is negotiated when the merchant signs the contract with the bank.[7] From the merchant's perspective, the transaction is now complete. However, the clearing and settlement process has only just begun.

Although the drafts are computer punch cards, they are not yet machine readable. Banks with very low volume may just manually sort and total the drafts, but others send them to the proofing and data entry departments to be manually key punched and proofed.[8] The drafts are then sorted by card number. On-us transactions are fed into the computer to update the cardholder accounts, and are then added by collation to the drafts already processed for each cardholder since the last billing cycle. At this time, most banks are still performing *country-club billing*, where the physical drafts are included with each statement.

All interchange drafts are then grouped and totaled by issuing bank. The merchant's bank

[6] At this time, interchange was rare on the average. However, there were localized exceptions to this. For example, the National Bank of Commerce in Seattle and Puget Sound National Bank in Tacoma experienced a high level of interchange due to the large amount of business that takes place between those two cities, which are roughly 30 miles apart. The authorization centers at these two banks simply called each other in the morning and kept a line open, allowing them to authorize interchange transactions quickly over a speakerphone (Jutilla interview).

[7] Merchant discounts at this time ranged anywhere from 0 to 8 percent, averaging 3.5 percent. Technically, some banks credited the merchant the full amount, and then debited the account for the discount fee at the end of the month (Jutilla, pp. 49, 179).

[8] *Proofing* involves verifying that the drafts total to the same amount claimed by the depositor. This was often done by encoding the human-readable elements of a draft into machine-readable form, so that the drafts can be machine-totaled.

completes a special *clearing draft* against the issuing bank for the total of all the sales drafts. The clearing draft looks very much like a cashier's cheque, complete with the magnetic ink routing characters, and can be submitted through the normal cheque clearing system for payment. The physical sales drafts on the other hand are mailed directly to the issuing bank through the US postal system. The clearing draft is often processed before the individual sales drafts arrive at the issuer. Thus the issuer is forced to transfer funds, but must wait until the sales drafts arrive to reconcile and add the charges to the relevant cardholders' accounts. Once they arrive, the issuer reconciles the sales drafts against the settlement payments, and then performs the same actions the original bank did for the on-us case.[9]

4.1.2 Operational Problems

Within this simple transaction scenario, we can begin to see a number of operational problems that were greatly exacerbated by the system's increasing sales volume.

Authorization, Floor Limits, and Fraud

The first notable operational problem was the interaction of authorization, floor limits, and fraud. Payment card transactions differ from those in other payment systems in one important way: they are *guaranteed*.[10] If the merchant follows the rules of the program, the merchant is guaranteed payment, even if the transaction was fraudulent. In the case of a cheque, the issuer simply returns the bad cheque and the merchant must absorb the loss; in a payment card transaction, the *issuer* must absorb the loss. This introduces a certain amount of risk to the issuing bank, and in an ideal world, the issuing bank would like to eliminate that risk by authorizing every transaction. However, this was not a realistic option in 1969, as the labor and telecommunication costs would easily outweigh the revenue gained from a low-value transaction. Additionally, authorizing every transaction would delay an already slow process, risking the use of cash or a cheque instead of the card.

The floor limit concept is essentially a cost/risk tradeoff made by the banks. Not all transactions are equally risky, and the easiest way to distinguish the higher-risk ones is by the combination of purchase amount and merchant type: a high-value purchase from a jewelry store is more risky than a low-value purchase from a shoe store. However, what most banks did not anticipate was that criminals would quickly discover the various floor limits and make numerous under-limit charges, resulting in significant losses. A new card stolen from a mailbox could be used for a week or more before the issuing bank even saw the first sales draft, and over a month before the cardholder received the first statement for a card the customer did not even

[9] In addition to merchant purchases, cardholders could also obtain cash advances directly from any BankAmericard bank. Although these transactions levied different fees and had no free period, they were processed in the same manner, with the bank providing the cash advance acting as the "merchant."

[10] Note that although cheques are not guaranteed, travellers cheques are.

know was issued.[11] Once detected, banks would notify other authorization centers and mail a postcard to merchants that might likely see the card.[12]

However, relying on the merchants to catch the cards was problematic. The main incentive for merchants to use the authorization system is the *guarantee of payment*, not the reward for catching a stolen card. The authorization process is more than just a technical function—it also formally transfers the responsibility for fraud from the merchant to the issuer. A merchant was (and still is) allowed to take a transaction above the floor limit without authorization, but the merchant then assumes the risk of fraud. If an issuer can prove that the merchant did not authorize the transaction, or that the bank warned the merchant about the card number prior to the transaction, the issuer can submit a *chargeback* into the system, which will eventually debit the merchant's account. However, proving a chargeback required a manual audit, and most bankcard processing centers were already struggling to keep up with the sharply-increased sales volume.

Merchants were also not inclined to call for authorizations due to the delay it would cause at the point of sale—sources from the time estimated that the average authorization took anywhere from five to twenty minutes, depending on how quickly the merchant could get through to the authorization center, and how quickly the merchant's bank could call or telex the issuing bank in an interchange case.[13] Stallwitz found that nearly all merchants in his study complained about the speed of authorization, and some admitted that they encouraged the use of cash or a cheque when the purchase was above their floor limit.[14] Others would rely on their own assessment of the customer (often based on appearance) and take the card without authorization, or simply reuse an authorization code from a prior transaction as it was unlikely that the issuing bank would detect this under the manual system of the time.[15] Stallwitz also found that suburban merchants in particular would avoid consulting the hot card lists and calling for authorization as it might offend their customers and risk the loss of the sale. Lastly, some merchants were themselves creating or participating in fraudulent transactions. Restaurant cashiers would make additional sales drafts with a customer's card, or less reputable merchants would submit under-limit drafts using a stolen card and split the proceeds with the thief.[16]

The actual amount of fraud occurring at this time is difficult to estimate as banks were not required to disclose such information, nor were they particularly eager to do so. Those that did were either inconsistent in the way they calculated and reported losses, or as Spencer

[11] Often the cards were actually stolen by the postal sorters and carriers. The practice of mailing unsolicited cards to consumers was eventually banned by the US Congress in 1970, and most other countries have since passed similar laws.

[12] Jutilla, p. 221-223. Eventually the Visa system produced a weekly booklet of hot card numbers, but this was ultimately replaced by online authorization via inexpensive point of sale dial terminals (see Chapter 8).

[13] "It took about 15 to 20 minutes to make a $35 purchase, which didn't make you very popular at the point of sale" (Russell interview). See also Stallwitz, pp. 44–45.

[14] Ibid., p. 45.

[15] Reusing authorization codes became much easier to detect after NBI computerized both authorization and clearing and settlement, the story of which will be told in the next two chapters.

[16] Jutilla, pp. 219–229; Nocera, p. 30; Galanoy, p. 149.

Nilson claims "doctored the records so that it would come out to a ratio acceptable to their peers."[17] Nevertheless, Nilson and others attempted to estimate how much the banks were losing on their card programs. The estimates range wildly and are difficult to compare as they are for different time periods, different sets of card programs (e.g., bankcards only, bank and T&E and retail, etc.), and different loss categories (total losses as opposed to losses specifically attributable to fraud). Nilson estimated that *fraud-specific* losses on bankcards increased from a mere $140,000 in 1967 to $2.2 million by 1969.[18] Various Federal Reserve studies reported that *total* losses for bankcards rose from $12 million in 1967 to $115.5 million in 1970.[19] Nocera claimed that throughout the late 1960s, the Chicago banks alone lost over $25 million, and the New York banks over $250 million.[20]

The growing amount of fraud was clearly a concern for those banks participating in the BankAmericard licensing program. Beyond the actual monetary losses, the shocking headlines were creating a *perception* that fraud was rampant and bankers were doing nothing to protect their cardholders.[21] This perception could not only erode the confidence of cardholders and merchants, but also attract the unwanted attention of lawmakers and regulators. Indeed the US Congress held hearings on the practice of mailing unsolicited cards in 1967 and was drafting legislation to not only prohibit it, but also protect consumers from the cost of fraudulent charges.[22] As is typical, these hearings became a thinly-veiled public trial of the entire bank credit card industry, accusing the banks of fueling inflation and tempting innocent consumers to abandon the traditional values of thrift in favor of reckless debt spending.[23]

[17] *The Nilson Report*, No 161, 11 Apr 1977. The general accuracy of the Nilson Report was contested by many of my interview sources, so some of his claims and statistics must be approached with caution.

[18] Ibid. Dollar amounts are in USD.

[19] 1967 data from 'Bank Credit-Card and Check-Credit Plans', Federal reserve system report (Federal Reserve System, July 1968); 1970 data reported in Phillip Brooke, 'Banks Reappraise Cards as Losses Mount', *American Banker*, 18 May 1971, p. 1.

[20] Nocera, p. 61.

[21] For example, see Galanoy. Formerly the Director of Communications for NBI, Galanoy accused bankers of being blinded by their desire to build an all-encompassing electronic funds transfer system, ignoring the costs of fraud to consumers. For an example of this concern voiced in the popular press, see Paul O'Neil, 'A Little Gift From Your Friendly Banker', *Life* 68 (27 March 1970).

[22] These laws were passed in 1970 as an amendment to the Truth in Lending Act (Roland E Brandel and Joseph E Terraciano, 'Legal and Regulatory Environment', in: *Critical Issues in Bank Cards* (Washington DC: American Bankers Association, 1980); Fisher et al., p. 257). The 1967 hearings are documented in *Unsolicited Bank Credit Cards*, Hearings before the Committee on Banking and Currency, United States House of Representatives, 19th Congress, First Session. (8 and 9 November 1967).

[23] Dee Hock provided perhaps the best rebuttal to this in a 1979 interview: "Sure, consumer debt is high, but if you want the consumer to stay out of debt, business and government have to set the example. If we expect consumers to reduce debt and increase savings, then we must create an environment without inflation and with tax laws that favor saving and not debt. After all, interest paid on debt is tax deductible and interest earned on savings is taxed. How can that encourage thrift? We should not criticize the consumer who is learning to play the government invented game of buying now through debt and paying later with inflated dollars." Bill Streeter, 'Let's Stop Using EFT Jargon, and Start Offering New Services', *ABA Banking Journal* 71:7 (July 1979), p. 75.

Clearing and Settlement of Interchange Transactions

The second major operational problem area was the clearing and settlement of interchange transactions. Like a cheque, a payment card sales draft is a claim on funds that must be cleared and settled with the issuing bank. If a different bank acquired that transaction, there must be a mechanism by which the draft can be routed to the issuer, and payment made to the acquirer.

As noted in the previous chapter, most banks at this time cleared and settled their cheques through the national clearinghouse operated by the Federal Reserve. It would seem that using this same system to clear and settle credit card sales drafts, which were small in number compared to cheques at this time, would be a sensible thing to do. The bankcard associations approached the Fed about processing credit card drafts, but the Fed refused to handle them.[24] Technically, it would have required some modifications to the automated systems: the sales drafts were 80 column IBM punch cards, larger in size than most cheques of the day; and they encoded information as punched holes instead of magnetic characters printed along the bottom edge. However, the technical reasons were secondary to the more ideological belief that debt instruments, especially those involving a discount, simply did not belong in the Federal Reserve's clearing system.[25]

With the Fed's refusal to handle credit card drafts, the BankAmericard Service Corporation (BASC) was faced with a problem: how should the licensee banks clear and settle their interchange transactions? One logical option would have been for the BASC to create their own centralized clearinghouse for BankAmericard transactions. However, the BASC chose not to do this, partly because the amount of interchange was still very low in the late 1960s.[26] Instead, the BASC stipulated that acquiring banks must mail interchange drafts directly to the issuing bank, similar to the way they handled out-of-town cheques in the nineteenth century. The issuing bank would then reimburse the acquiring bank, less a discount fee, called the *interchange reimbursement fee*.[27]

However, this solved only the clearing half of the problem—the licensee banks still needed a way to settle those transactions (i.e., transfer "good and final funds" from the issuer to the acquirer). Recall that the Federal Reserve System eliminated the need to transfer physical currency between banks when settling payment transactions, and the BASC decided to leverage this system by creating a special *clearing draft*, which looked like a bit like a cashier's cheque.

[24] Russell interview. Hock commented on this in a 1974 speech: "Had the Federal Reserve agreed when asked (and they were) to clear bank card activity, would the service have evolved as it subsequently has?... It is clear there would be no BASE II and no INAS today had the Federal Reserve said yes, and clear that present bank card service would be radically different" (Dee Hock, 'Electronic Funds Transfer or Electronic Value Exchange?' (October 1974), p. 21).

[25] Russell interview.

[26] Sources estimated that it was between one and five percent of transactions at the most. There were of course localized exceptions to this. In regions where banks were not allowed to operate branches across an entire metropolitan area, the interchange level would naturally be higher.

[27] Note that the laws governing cheque clearing discussed in the previous chapter did not apply to credit card sales drafts. Any similarity in their clearing method was coincidental and not required by law. The legal basis for credit card sales drafts came from the contracts signed by licensee banks, cardholders, and merchants (Katz interview).

To receive payment for a set of interchange drafts, acquirers completed one of these clearing drafts against the issuer for the total amount of the sales drafts, less interchange fees, and submitted it along with their other inter-bank funds transfer requests.

This separation of the clearing draft from the sales drafts allowed banks to use their existing funds transfer mechanisms, but it also created a timing problem that jeopardized the functioning of the entire system. When the issuing bank received payment notice of the clearing draft, it would enter that amount into a suspense ledger and wait for the individual sales drafts to arrive in order to reconcile and bill the cardholder. Unfortunately, this often took quite a long time. This is how Visa's founder described it:

> Meanwhile, the merchant bank, having already been paid and under immense pressure to handle its own cardholder transactions, had no incentive to process foreign transactions and get them to the issuing bank for billing to the cardholder. Since each bank was both a merchant-signing bank and a card-issuing bank, they began to play tit-for-tat, while back rooms filled with unprocessed transactions, customers went unbilled, and suspense ledgers swelled like a hammered thumb. It became an accounting nightmare.[28]

This immense backlog in the system also compounded the fraud problems discussed earlier. Issuing banks would have no way of knowing if sub-floor-limit fraud was occurring on a card until the actual sales drafts arrived and were processed. By the time they arrived, thousands of dollars worth of fraud could have taken place.

Even when the sales drafts did arrive, it was often the case that their total did not match the clearing draft amount. Many smaller merchant banks would simply run an adding machine tape over the drafts instead of key punching them, and would inevitably make mistakes. Chuck Russell, who succeeded Hock as CEO, recalled that "Banks couldn't balance from day to day because they couldn't get their drafts drawn on other banks settled. It was a disaster."[29] To provide a sense of the scale of the problem, he relayed this story:

> I was shown a room that was warehoused-sized, full of IBM 80-column tab cards (which were the drafts) that they couldn't settle. We're talking millions and millions of dollars...they had never got the debit or the credit side of the transaction through clearing because they couldn't find them![30]

Finally, it should be noted that not all banks experienced problems to the degree described here. However, the lack of a centralized clearinghouse, compounded with the timing problems introduced by the clearing drafts, created operational problems that were most definitely threatening the overall system's stability.[31]

[28] Hock, *One From Many*, p. 77. Note that by "foreign," Hock simply means interchange transactions, not necessarily international transactions.

[29] Russell interview.

[30] Russell interview. The "stacks of unprocessed drafts" story was also relayed by others in various forms.

[31] Jutilla indicated that his bank was typically able to reconcile, but the delays in receiving the interchange drafts were especially dangerous due to fraud. He concluded that the system could not have survived the way it was as the transaction volume increased.

4.1.3 Organizational Problems

Although the operational problems described in the previous section may have had potential solutions within the existing franchising organization, that organization had problems of its own that further compounded the operational difficulties. I will now briefly review these organizational issues to show that in order for the system to survive, a new type of organization was necessary.

The BankAmericard licensing system, like any cooperative payment system, faced a central organizational tension—balancing competition and cooperation. The licensing system created a new meta-organization comprised of *competing* financial institutions that needed to *cooperate*, at least to some degree, in order to provide a universal payment system that none could have realistically provided alone. Competing organizations in a marketplace normally seek their own self-interests in an assumed zero-sum game for market share. A cooperative organization, on the other hand, offers a different possibility—if all members cooperate, they can provide a larger, universal system that allows them all to benefit even more than if they chose not to cooperate. In other words, each participant's slice of the cooperatively baked pie would likely be larger than any pie the participant could have baked alone. However, to accomplish this, they need mechanisms that will create trust within the organization, mechanisms that balance out their power and interests and dictate how inter-organizational work will be accomplished. In other words, they need something akin to a constitution, as well as *operating regulations*, to which all member organizations agree.

Under the BankAmericard licensing system, BofA retained not only the ownership of the BankAmericard name and marks, but also all the power, and this led to a fundamental distrust between BofA and the licensees. The licensees knew that BofA would have opened branches in their territories if the banking regulations had allowed it, and if those regulations ever changed, BofA could easily revoke their license and become the sole BankAmericard issuer.[32] The licensees also doubted if BofA had the desire and even the ability to solve the operational problems discussed earlier.[33] The licensees believed that any solutions developed by BofA would naturally be in BofA's best interest and not those of the licensee banks.

Although BofA retained nearly all the power in the system, their power to enforce and modify the operating regulations was neutered by two critical flaws in the license contracts. First, the contracts lacked mechanisms for financially punishing banks that skirted or bent the operating regulations, nor did they contain a method for resolving grievances between the licensee banks. The only recourse BofA had was to revoke a bank's license, but since most of these banks held large correspondent deposits with the BofA, and were dominant in their geographic area, this was not likely to happen. Second, the contracts also lacked a clause

[32] Hock, *One From Many*, p. 85. Of course, these regulations were abolished in the 1980s, but by then it was too late, as the Visa system had already been established.

[33] Russell interview. The BofA paid very low salaries at the time, and the most talented operational people tended to go to their main competitor, Wells Fargo, which was a member of the Interbank system.

allowing BofA to change the operating regulations in response to new developments. If BofA needed to modify or add a rule, they had to re-negotiate a new contract. Again, BofA had no recourse if banks simply refused to sign the new license, which they often did if the rules were not in their best interests.[34]

The fundamental distrust and the flaws in the contracts created a number of organizational instabilities. The most significant and pernicious was the tension over the interchange reimbursement fee. As noted earlier, this fee was paid by the acquirer to the issuer during the settlement of an interchange transaction.[35] At this time, the intent of the fee was to compensate the issuer for the cost and risk of extending the cardholder credit for the transaction. The rule established under the licensing system for interchange fees was essentially unenforceable. This is how Bennett Katz, Visa's long-time general counsel described it:

> When I came on board, the rule was…if a customer of your bank goes into a merchant belonging to another bank, outside of that territory, then the bank that signed the merchant has a choice as to what it sends to the issuer. It could send the amount of the discount that it received from the merchant less a processing fee (for processing the transaction), or if it didn't want to calculate each and every one…it could send the average discount it was getting from all of its merchants less a processing fee. Well they would say 'my average is two percent.' How are you going to audit that? And if the merchant put up a big deposit, their merchant discount might be close to zero, and the issuer would get almost nothing!
>
> So the issuer has all the costs because he's extending the credit and eating defaults, but he was getting almost nothing when the customer traveled. The losses were horrendous. It was literally chaos in the BankAmericard system.[36]

4.1.4 Tensions Come to a Head

In October of 1968, the BASC called a special meeting of the licensees to discuss the operational and organizational problems facing the BankAmericard system. Card program managers from each of the licensee banks descended on Columbus, Ohio, but the BASC neglected to send their most senior officers. The licensees were incensed that the BASC apparently did not recognize the seriousness of the situation, and began to make accusations that the BASC was either unwilling or incapable of solving the system's problems. By the middle of the second day, the meeting had devolved into "acrimonious argument."[37] Unsure of how to rescue the situation, the BASC representatives attempted to create a committee of licensees that would look into the most critical problems. However, one of those selected to be on the committee had a different idea of what it would take to solve the system's problems, and after lunch the

[34] Katz interview. See also Hock, *One From Many*, pp. 83–87.

[35] The acquirer paid this out of the discount fee collected from the merchant in return for immediate and guaranteed payment. However, in the case of a cash advance, the flow is reversed: the issuer pays the acquirer, because the acquirer assumes all the costs.

[36] Katz interview.

[37] Hock, *One From Many*, p. 84.

rest of the licensees were greeted by the card-center manager from the Seattle National Bank of Commerce: Dee Ward Hock.

4.2 Dee Hock

Describing Dee Hock is a formidable task, but one that must be attempted since the design of the Visa organization is very much a reflection of him.[38] He is an extremely complex person who defies any attempt at categorization. Those that worked at Visa during Hock's reign would invariably start telling "Dee stories" during our interviews. However, no single story could entirely capture his character, so they just kept telling more of them. Tom Cleveland, who eventually became Visa's CFO, wrote a series of 16 stories totaling 50 pages to try to capture what he called "the Dee Hock experience."[39] Sources used a wide variety of adjectives, but none of them seemed to be sufficient: inspirational; intimidating; clairvoyant; clever; shrewd; aggravating; fair; brutal; demanding; and brilliantly eccentric. Many said he was the most decent human being they had ever met, and most said he was nearly impossible to describe fully.

Interestingly, all of the stories often combined to describe Hock as simultaneously having one character trait and its opposite. He was the most inspirational of leaders, but could also be the most denigrating. One employee said that he was like an "emotional roller-coaster"—he could completely demoralize you in one breath and then boost your ego to new heights in the next. Hock advocated decentralized autonomy for his employees and gave them enormous leeway at times, but would ultimately micromanage every detail he could. Every piece of correspondence that left the Visa offices during the first few years was personally reviewed by Hock, and he reportedly obsessed about the design of each new office or data center space, even down to the location of the electrical sockets. In short, his personality is exceedingly complex, and often paradoxical, which makes him quite difficult to describe. Indeed he might be best described as complexity personified.

4.2.1 Early Life

Despite the complexity of his personality, his own story is almost the stereotype of the American, pioneering, self-made man. Born in 1929 at the beginning of the Great Depression, Hock grew up in a poor family in rural Utah. He worked a number of harsh manual-labor jobs in his youth, which is what "proud men did, without whining. 'Root, hog, or die' was the homily of the day."[40] Hock was also a curious child, a voracious reader, and a keen study of nature.

[38] The most complete profile of Hock is his own autobiography (Hock, *One From Many*, or the previous edition, Hock, *Chaordic Age*). See also Chutkow, pp. 81–98; Nocera, pp. 62–66; and Dougherty.

[39] Tom Cleveland, 'The Visa History: Tom Cleveland's Perspective' (April 1999), p. 1.

[40] Hock, *One From Many*, p. 24. Hock's autobiography puts forth this pioneering, self-made persona, and to a large degree, it was confirmed by interview sources who knew him in the 1970s.

In High School he joined the debate team where he learned how to argue persuasively. He went on to be the first person in his family to attend junior college where he was introduced to philosophy and literature, which eventually shaped his views on organizations and leadership. Although he earned an Associate's degree, Hock considered himself to be mostly self-taught, and throughout his life he retained the idea that too much formal schooling was not only a sign of undue privilege, but also a detriment to the lessons taught by the world's best teacher: experience.

Hock was especially hostile to MBAs, as he thought that business schools "essentially teach mismanagement."[41] Consultants with MBAs were anathema, but a necessary evil at times, as they lent an important credibility to Hock's ideas. However, as one of his executives explained, Hock always made the MBA's realize that they still had much to learn:

> He hated MBAs…he hated the whole group of MBAs as a general rule…. He used to ask me "what's your most important job here at Visa?" Well I would give him the standard MBA school of thought at the time. And he would say "bullshit! That's your problem."—and he told me this probably a hundred times—"you're primary objective is to make me look as good as I possibly can." But he's right![42]

Ironically, a large number of Hock's key employees actually did hold MBAs and a few earned PhDs, but they were typically counseled to keep these accolades to themselves.

Although Hock had a dim view of formal education, this is not to say that he is anti-intellectual. He is, in the very literal sense, a philosopher, and his sheer brilliance is uncontested. Although most of those who knew him acknowledged that he had a difficult and abrasive personality, they also admitted that he was the smartest person they had ever met. The word most commonly used by my interview subjects when describing Hock was "genius." Tom Cleveland remarked that "If ever there was evidence that aliens added to our gene pool, Dee is it…. He is cut from another cloth, definitely not from this world. I have yet to meet anyone so empowered on this planet."[43] Frank Fojtik, one of Visa's key system engineers, put it this way: "Most human beings, no matter how hard they try, are restricted in their way of thinking. He's not."[44] Hock's intelligence, combined with his fine-tuned rhetorical skills, allowed him to both inspire those who followed him, and outmaneuver most of those that resisted him.[45]

4.2.2 Views on Organizations

Hock's first professional job was with a consumer finance company in Los Angeles starting in 1951, and it was here that he not only learned the business of lending, but also began to develop

[41] Dougherty, p. 14.
[42] Anonymous interview.
[43] Cleveland interview.
[44] Fojtik interview.
[45] Hock made this point himself when he wrote "During my years of college debate, I had held fast to the notion that until someone has repeatedly said 'no!' and adamantly refuses another word on the subject, they are in process of saying 'yes' and don't know it" (Hock, *One From Many*, p. 119).

his concepts about organizations. During this time, Hock became increasingly suspicious and critical of what he called "mechanistic, command-and-control organizations."[46] These types of organizations are typified by the centralization of power, the creation of bureaucratic hierarchies, and the use of technology and highly-rationalized rules to control an increasingly specialized and deskilled set of workers.[47] For Hock, this concept of organizational design was bred from "industrial age thinking" and the "machine metaphor," where an organization is viewed as a sort of machine with humans as the cogs and wheels.[48] Hock sees these organizations as fundamentally flawed—in these kinds of organizations, "purpose slowly erodes into process," "procedure takes precedence over product," and "the doing of the doing" causes nothing of substance to get done.[49] Furthermore, Hock argues that command-and-control organizations are actually incapable of dealing with their increasingly complex and dynamic challenges, and because this is still the dominant organizational form, we are in "the midst of a global epidemic of institutional failure."[50]

Hock would eventually develop a different concept for organizations based on his observations of natural systems.[51] He created the neologism "chaordic" to describe his new concept and defined it as "The behavior of any self-organizing and self-governing organism, organization, or system that harmoniously blends characteristics of chaos and order."[52] In contrast to the command and control organizations, chaordic organizations are decentralized, self-organizing, self-governing, exhibiting emergent properties. The obvious implication is that they are also more flexible and adaptive.[53] However, this concept was still rather nascent in Hock's mind at this time, and we shall see how he continues to develop it through the creation of the Visa organization, which he considered to be the first, albeit flawed, implementation of this new concept.[54]

Hock began to implement some of these new organizational ideas while working for the consumer finance company in Los Angeles. Unfortunately for Hock, that organization did not share his vision, and eventually fired him. It was at this point that Hock learned his most important lessons about consumer lending. While he was still employed, he had gotten into "considerable debt" through the use of credit cards, and had no savings to live upon. Hock took multiple jobs to pay it off, and he and his wife vowed never to carry so much debt again. He

[46] Hock, *One From Many*, p. 36.

[47] Hock purposely does not cite his sources, but this is of course reminiscent of Burns and Stalker's 'mechanistic' and 'organic' organizational spectrum. See Tom Burns and G M Stalker, *The Management of Innovation* (London: Tavistock, 1961). One can also detect resonances with General System theory, Chaos theory, and what is now begin called "ecological thinking."

[48] Hock, *One From Many*, pp. 37–38.

[49] Ibid., p. 36.

[50] Ibid., p. 11.

[51] Ibid., pp. 248–249.

[52] Ibid., p. 13.

[53] Burns and Stalker argue that this organic style of organization is more effective for firms in highly-dynamic markets, while the mechanistic style is best for stable markets. However, Hock would question if any organization exists in a stable context, and point out that mechanistic organizations will ultimately fail due to their inherent flaws.

[54] Ibid., pp. 248–249.

wrote, "It is amusing now to remember how we shredded every credit card in our possession, swearing never to have another."[55]

Hock moved to Seattle and interviewed for a job at the National Bank of Commerce in 1965.[56] He played several menial roles within the bank before being picked in 1966 to help start their newly-licensed BankAmericard program. He and his boss were given a ridiculously short time in which to start the program—90 days—and by the end of their training session in San Francisco, they realized that they were in "deep trouble."[57] Hock discovered that there was "bad blood" between the BASC and BofA's own card operations center, and that the card center "had no capacity to comply with rules prescribed in the licensing agreements."[58] The training, provided by the marketing department, amounted to encouragements to mass-issue the cards, with no practical advice on how to handle the aftermath. Hock and his boss spent the next week visiting every card center they could in order to learn the realities of credit card operations. Hock summarized: "Within a single week, our original belief that the BankAmericard franchise would provide a well-marked, expeditious road to the future had been shattered by what we had learned."[59]

The highly dynamic, unpredictable nature of the endeavor provided Hock with another chance to try out some of his new organizational concepts. He suggested that they "abandon tradition, throw detailed planning to the winds, rely on a clear sense of direction, a few simple principles, common sense, trust in the ingenuity of the people, and let the answers emerge."[60] Despite numerous setbacks, his new organizational style proved effective, and with great effort they issued 100,000 cards and signed up enough merchants, all within 90 days. Throughout the process, Hock had become a believer in the future of payment cards, but not because of their financing capabilities. Hock wrote that while he was organizing the new program, he was also reading about the possibilities of a more complete integration of computers and banking. "Giving people another way to borrow money interests me not at all," Hock recalled, "What credit cards might become is something else again."[61]

4.2.3 Views on the Nature of Money

Hock had been slowly coming to the realization that "money" had become nothing more than "guaranteed alphanumeric data" and that a bank is nothing more than an "institution for the custody, loan, and exchange" of this data. Furthermore, that data was increasingly being stored and manipulated by computers, and would eventually "move around the world at the speed of

[55] Hock, *One From Many*, p. 40.
[56] Dougherty, p. 13.
[57] Hock, *One From Many*, p. 61.
[58] Ibid.
[59] Ibid.
[60] Ibid., p. 62.
[61] Ibid., p. 69. Although these quotes come from a retrospective biography, his speeches from the 1970s contain similar sentiments.

light at miniscule cost by infinitely diverse paths."[62] He then came to one of his most important conclusions:

> Any institution that could move, manipulate, and guarantee alphanumeric data in the form of arranged energy in a manner that individuals customarily used and relied upon as a measure of equivalent value and medium of exchange was a bank. It went even beyond that. Inherent in all this might be the genesis of a new form of global currency.[63]

Hock realized that if this was actually the case, the implications were enormous:

> If electronic technology continued to advance, and that seemed certain, two-hundred year old banking oligopolies controlling the custody, loan, and exchange of money would be irrecoverably shattered. Nation-state monopolies on the issue and control of currency would erode.... The vast preponderance of the system would fall to those who were most adept at handling and guaranteeing alphanumeric value data in the form of arranged particles of energy.[64]

Lastly, Hock realized that he, and most of his fellow bankcard managers, had misunderstood what business they were in:

> It seems ordinary and obvious now. It was a revelation then. We were not in the credit card business. "Credit card" was a misnomer based on banking jargon. The card was no more than a device bearing symbols for the exchange of monetary value. That it took the form of a piece of plastic was nothing but an accident of time and circumstance. *We were really in the business of the exchange of monetary value.*[65]

4.3 Creation of National BankAmericard Inc.

At the licensees meeting in 1968, Hock convinced the BASC representatives that merely investigating a few of the most serious operational problems would do nothing to save the system. Instead he proposed a "cohesive, coherent, self-organizing effort involving all licensees to examine all problems plaguing the system."[66] The BASC representatives were naturally reluctant to agree, but they had little reason to object—Hock's proposal did not require them to commit to any policy recommendations made by the licensees, nor could the BASC prohibit the licensees from meeting on their own and making plans that might exclude the BofA.

Hock created a matrix of committees to investigate the problems of the system. He divided the United States into eight regions and each region formed separate committees to focus on four functional areas: operations, marketing, credit, and computer systems. After performing some initial research, the committees began to report:

[62] Hock, *One From Many*, pp. 95–96.
[63] Ibid., p. 96.
[64] Ibid., p. 97.
[65] Ibid., p. 98. Emphasis in original.
[66] Ibid., p. 87.

> The complex of committees had but one redeeming quality: It allowed organized information about problems to emerge. It took only two cycles of meetings to realize that the problems were enormously greater than anyone imagined—far beyond any possibility of correction by the existing committees or the licensing structure—and growing at an astonishing rate. Losses were not in the tens of millions, as everyone had thought, but in the hundreds of millions and accelerating.[67]

It quickly became obvious to Hock that the licensing system could not survive as it was structured. The problem was not a lack of technology—it was a *fundamental flaw in the organizational design*. If the licensing organization could barely function in the United States and a few other countries, it had no chance of expanding into the ubiquitous worldwide system that Hock envisioned. His new organizational ideas had been successful in the highly dynamic and chaotic context of starting a new card program at a particular bank, but could they also work to organize a system that involved hundreds, and eventually tens of thousands, of independent organizations spread across different countries, subject to different regulations? If so, there was an enormous opportunity:

> Any organization that could guarantee, transport, and settle transactions in the form of arranged electronic particles twenty-four hours a day, seven days a week, around the globe, would have a market—every exchange of value in the world—that beggared the imagination. The necessary technology had been discovered and would be available in geometrically increasing abundance at geometrically diminishing cost. But there was a problem. No bank could do it. No hierarchical, stock corporation could do it. No nation-state could do it. In fact, no existing form of organization we could think of could do it. On a hunch I made an estimate of the financial resources of all the banks in the world. It dwarfed the resources of most nations. Jointly they could do it, but how? It would require a transcendental organization linking together in wholly new ways an unimaginable complex of diverse institutions and individuals.[68]

Hock selected three members of the national executive licensee committee, which had overseen the work performed by the regional committees, to help him think through a design for a new organization. They checked into the Alta Mira hotel in Sausalito, California for a week of discussion. The location was perhaps fitting—it is perched upon a hill across the bay from San Francisco's financial district, home of the BofA. It was here that Hock and his colleges, "the founding fathers," planned what amounted to a revolution against the BofA's licensing system, though the BofA itself would continue to play an important role in the new system.[69]

Hock recounts that on the fourth night he began to realize that if the methods of biological evolution could produce such complex organizations as brains and immune systems, not to

[67] Hock, *One From Many*, p. 91.

[68] Ibid., pp. 98–99.

[69] Nocera, p. 91. Indeed when Hock and his comrades finally met with Ken Larkin, the Vice President in charge of the BankAmericard, Larkin greeted them by saying "Here they are—the leaders of the revolution!" (Hock, *One From Many*, p. 116).

mention larger and more diverse systems such as rain forests, marine and weather systems, the same sort of principles employed by humans might be able to create the kind of complex, self-organizing system he wanted to achieve. He concluded, "What if we quit arguing about the structure of a new institution and tried to think of it as having some sort of genetic code?"[70] He continued, "If institutions have no reality save in the mind, might their genetic code have something to do with beliefs—with purpose and principles?"[71] What if they came to agreement on the purpose and principles of the new organization, and then let it self-organize according to those?

Hock presented his ideas the next morning and the group formed the following principles, which would in turn become the basis for the new organization's constitution:

> *What if ownership was in the form of irrevocable rights of participation*, rather than stock: rights could not be raided, traded or sold, but only acquired by application and acceptance of membership?
>
> *What if it were self-organizing*, with participants having the right to self-organize at any time, for any reason, at any scale with irrevocable rights of participation in governance at any greater scale?
>
> *What if power and function were distributive*, with no power vested in or function performed by any part that could reasonably be exercised by any more peripheral part?
>
> *What if governance was distributive*, with no individual institution or combination of either or both, particularly management, able to dominate deliberations or control decisions at any scale?
>
> *What if it could seamlessly blend cooperation and competition*, with all parts free to compete in unique, independent ways, yet able to yield self-interest and cooperate when necessary to the good of the whole?
>
> *What if it were infinitely malleable, yet extremely durable*, with all parts capable of constant, self-generated, modification of form or function without sacrificing its essential purpose, nature, or embodied principle, thus releasing human ingenuity and spirit?[72]

Hock persuaded the other licensees to agree to these principles, but as he admits, this may have had more to do with their beliefs that the entire project was likely doomed to failure anyway. Hock recalled:

> In the beginning, none of the licensees thought that the Bank of America would surrender ownership of a trademark and licensing system that assured them a quarter percent or more of the revenues of every participant in perpetuity. No one

[70] Hock, *One From Many*, p. 108.

[71] Ibid., p. 109.

[72] Ibid., pp. 109–110. Emphasis in original. It should be noted that the Interbank organization was also founded on many of these same principles, though perhaps for more practical reasons (Honey interview). Although Hock wanted to balance malleability and durability, the former seems to have won out over the latter. Visa recently announced it would restructure into a public stock-issuing corporation, essentially abandoning these principles.

> thought that banks would voluntarily surrender a portion of their autonomy to an external entity in order to act together for a common purpose. No one believed that such a horizontal grouping of competitors could exist within the spirit and constraints of antitrust laws. And no one dreamed the emerging ideas would bring together in common ownership and enterprise people and institutions of every race, language, custom, and culture—every economic, legal, philosophical, and religious persuasion in the world.[73]

Hock then presented the principles and a rough idea of the new organization to Ken Larkin, BofA's Senior Vice President in charge of the BankAmericard program. Larkin's response was less than encouraging and the BofA initially resisted any attempts to curtail their power or control over the system. However, BofA actually had little choice but to capitulate and allow Hock to pursue the new structure. The system was clearly falling apart, and there was a very real danger that the licensees could organize a new competing system on their own and exclude the BofA, or simply join the rival Interbank network. Hock used his talents of persuasion to argue that it was actually in BofA's best interest to support the new organization, as it had the potential of generating far more revenue for them in the long run, and eventually they agreed to do so.

Hock began crafting the organization's structure, bylaws, and operating regulations. It would be a membership cooperative where the members would also be the joint-owners. In essence, the customers and owners would be the same, so there would be no divided loyalty. It would be non-stock, and membership would be non-transferrable, so that ownership and participation would forever be linked.[74] Membership qualifications would be set by the Board, but then any organization meeting those criteria must be allowed membership. Both membership fees and voting rights were determined by the amount of volume a member generated, linking taxation and representation. Operating regulations dictated not only how inter-member work was to be accomplished, but also how members would be penalized when they violated them. But most importantly, members would agree to abide by a common set of bylaws and operating regulations "as they now exist or are hereafter modified."[75] Thus Hock ensured not only that the new organization could enforce the rules, but also that they could modify and extend them as needed without renegotiating contracts.

Interestingly, Hock claims that he always wanted to include participating merchants as full owners and members in the new organization, but the idea was always strongly resisted by the licensee banks.[76] At the time, the licensee banks were concerned that giving merchants voting rights would endanger the merchant discount, which was a key source of revenue for acquirers. Hock also claims he wanted to include all the cardholders as owner/members, but that would have added a layer of complexity and chaos comparable to a democratic nation-state,

[73] Hock, *One From Many*, p. 112.
[74] Unfortunately, forever only lasted 36 years. See note 72 earlier in this chapter.
[75] Ibid., p. 124.
[76] Ibid., p. 161.

and perhaps would have been unrealistic.

The bylaws acted like a kind of constitution for the new organization. They enshrined Hock's principles, and although they could be amended by the Board of Directors, doing so required an 80 percent majority vote. This ensured that the organization could adapt if the principles no longer made sense, but they also could not abandon them without an overwhelming consensus.

The Board of Directors for the new organization was also designed to balance the interests and power of all the members, regardless of size and location, and to provide the larger banks with "blocking power" on key issues such as budgets and fees.[77] Each Board member would come from a different bank, so that no one bank could dominate policy regardless of volume. Each of the eight regions would be allocated one Director in order to ensure that all regions were equally represented. Five "at large" Directors would be elected by the general membership. One Director would be elected by the smallest banks in the system. Any bank that generated more than 15 percent of the sales volume would be able to appoint their own Director.[78] The president of the new organization (which would be Hock) would also have a seat on the Board, but could not serve as chairperson.

The functional committees would also remain in the new organization, though they would be split into two levels. Hock felt that it was important to keep the Board members occupied with the large and important policy decisions, while sub-committees handled the matters of minutia.[79] Thus he created a lower layer of "advisor groups," which were comprised of card-center managers. These groups reviewed policy proposals made by the central organization's staff, and occasionally drafted some of their own. The advisor groups then made recommendations to the Board-level committees, which were comprised of Directors, plus a member of the central organization's staff. The Board-level committees then voted whether to bring the policy before the entire Board. However, certain committees, such as the one for operations, were eventually given the power to approve changes to the operating regulations on their own without a full Board vote, as these changes were numerous.

In practice, the committees and advisor groups also provided the staff of the new organization with a testing ground for innovative ideas, enabling them "take the temperature" of the membership on important issues.[80] The staff could thus determine which issues would become politically contentious, which required some alteration, and which could be proposed with confidence at the full Board meetings. The relationship was symbiotic, as the committee and group members often became the champions of certain issues, enabling NBI's staff to appear as neutral advocates for the needs of the system as a whole.

The complex structure of the Board and committees was certainly a political masterpiece,

[77] Honey interview.
[78] This was eventually reduced to five percent (Honey interview).
[79] Honey interview.
[80] Honey interview.

and in principle, it ensured that while groups had blocking power, no one bank or person could dominate the organization. However, in practice the design did allow one person in particular to exert enormous influence: Dee Hock. By stipulating that the president of the new organization could not chair the Board, Hock had limited his own power to that of persuasion, but as Nocera remarked, "this was Hock at his most coy. It was precisely his ability to persuade that made him powerful."[81] Chuck Russell, who would succeed Hock as CEO in 1984, noted that Hock also used the structure of the Board and committees to his advantage:

> ... [Hock] structured the Board in such a way that *he* ran the company and there was no question about it. What he did, he very cleverly split up the power of the big banks into different groups. And if you get a Board that is widely split, *management* runs the company. Dee understood that, he understood it very well, and he taught it to me very well, because when I took over, man I kept that thing split into smithereens. You want a Board of at least 25 guys, because they'll never accomplish anything. Then you want to form lots of committees to supposedly do something, but you very strategically put fighting banks on each committee.[82]

A number of sources remarked that during the first decade of the Visa organization, it was in essence a dictatorship, and the only true power-broker was Hock. This is in part due to Hock's early successes in automating the core system processes, which will be discussed in detail in the next two chapters. In the minds of the member banks, Hock literally *saved the system* as well as their card programs, turning them from a source of major losses into a profit-making department. This gave Hock a large amount of influence with the Board, and its structure further ensured that he could shape the system towards his own vision: "the world's premier system for the exchange of value."[83]

Convincing the 200 fully licensed banks, and their thousands of sub-licensees, to join the new organization was no easy task, but Hock was persuasive and persistent. In 1970, only 90 days after Hock began the formation process, all the licensees had agreed to join, and the new organization was legally formed as a Delaware membership corporation named National BankAmericard Incorporated (NBI). The BofA became just another member, though they retained five special seats on the Board for the first few years to recognize their unique contribution in forming the original system.[84]

The exact legal description of the organization is a "for-profit, non-stock membership corporation."[85] This rather unusual classification is actually quite important. Banks at this time were not allowed to own stock in anything but a Bank Service Corporation (BSC), but BSCs were also subject to stringent regulation. By creating NBI as a non-stock membership cor-

[81] Nocera, p. 92.

[82] Russell interview.

[83] According to Nocera, Hock actually convinced the Board to adopt a resolution with that exact wording, even though most of the members had no idea what he meant by that.

[84] Nocera, p. 92. These seats were reduced by one each year until Ken Larkin was the only remaining director from the BofA.

[85] Katz interview.

poration, Hock enabled the banks to "own" it through membership, but avoided government regulation. Although NBI is legally formed as a for-proift enterprise, its owners are its customers, so it effectively operates as a non-profit entity. Because it does not issue any stock, and therefore has no market capitalization nor pays any dividends, any "accumulated net revenue" from member fees is used to finance the ongoing work of the organization.

4.4 First Steps Towards Building a Cohesive Payment System

With the organization now formed, Hock set about the task of building the cohesive, ubiquitous payment system he envisioned. In this section we will review some of the first steps taken by Hock and the new organization. These steps sought not only to solve a few of the most pressing problems, but also to reorient the thinking of the various member banks toward the good of *the system as a whole*, and not just the good of their individual programs.

4.4.1 Creating the Operating Regulations

Hock's first task was to establish a set of operating regulations. The operating regulations ("op regs" in Visa-speak), are essentially the rules of the game. They stipulate not only card design standards and rules for how the marks can be used in advertising, they also dictate how the various inter-member work should be accomplished and what penalties a member suffers if that member violates the rules.[86] The operating regulations also stipulate the basic rules an acquirer must enforce in their contracts with merchants. They are the central coordination mechanism for the entire system, providing a structure in which the competing members can cooperate.[87]

However, the operating regulations were written as a set of rules, and as such, they were in principle always open for interpretation. As the system developed and grew, disputes over transactions and chargebacks often called into question exactly what some of the regulations meant. To resolve these disputes, those in charge of maintaining the operating regulations also served as a kind of judiciary, arbitrating between the member banks, and clarifying the relevant regulations.[88] I will return to this rather interesting function in Chapter 10 when I discuss the need for such a judiciary in any cooperative network of competitors.

4.4.2 Building a Staff

Hock's second task was to hire a staff that could help him achieve his vision. Hock was an idealistic visionary, but he lacked many of the skills needed to run a day-to-day operation.

[86] According to Russell, the original op regs were almost a verbatim copy of those used by Interbank, but they quickly began to change and expand as the system developed (Russell interview).

[87] In Gerson's terms, the operation regulations are a type of *bracket* that facilitates inter-organizational work (Elihu M Gerson, 'Reach, Bracket, and the Limits of Rationalized Coordination: Some Challenges for CSCW').

[88] Tindal interview.

Shortly after NBI was formed, Hock hired a bank operations expert named Chuck Russell.[89] Russell had quite a bit of experience with bank credit cards: he started Pittsburgh National Bank's private card program in 1965, was a founding member of Interbank, serving on their operations committee, and later established Wachovia's Interbank card program in 1968. Russell complimented Hock's abilities and personality in a few important ways. First, Russell was a practical bank operations person who could put Hock's creative vision into action. Second, although Hock had spent some time at a bank, his roots were much more in finance companies and consumer lending, and Russell's more extensive banking background allowed him to temper Hock's enthusiasm with the realities of a more traditional banker's perspective. Lastly, while Hock could easily be abrasive and stubborn, the member bankers found Russell much more approachable and helpful.[90]

Russell and Hock created a powerful team, and both of them were necessary to build and manage NBI during its early years. Russell noted:

> We covered each others' blind sides beautifully.... If I was Marshal and I was trying to figure out how to invade Europe, Dee would be my guy. If I wanted someone to lead the charge, it sure as hell wouldn't be Dee! He'd be thinking three moves down the board and get killed in the process.[91]

Russell would later succeed Hock as CEO in 1984, retiring from Visa in 1994.[92]

Of course, Russell was not the only key figure hired during these first two years, and many others were also necessary to make the new system successful. Bob Miller, one of Hock's colleagues at the National Bank of Commerce, joined the effort shortly after NBI was formed. Bennett Katz, Visa's long-time chief counsel became the eighteenth employee in 1970. Ron Schmidt, the financial architect behind the new interchange fee, quarterly profit analysis and functional cost study, joined at the beginning of 1971. Aram Tootelian, Dave Huemer and David Goldsmith were hired in 1972 to lead the development of NBI's computerized authorization system, which I will discuss in detail in the next chapter. Bob Sanders was hired to develop NBI's adversing and marketing campaigns, but as discussed later in this chapter, Hock retained tight control over this area.

4.4.3 Fixing the Interchange Reimbursement Fee

Hock's next step was to fix the interchange reimbursement fee problem. Recall that this fee was paid by the acquirer to the issuer during the settlement of an interchange transaction. Under the licensing system, the rule was quite vague and acquirers often sought ways to avoid paying

[89] The initial staff hired before Russell actually planned a "palace revolt" against Hock, who they considered to be unreasonable, and were summarily fired (Cleveland, pp. 2–3; Russell interview).

[90] Ken Larkin once commented, "When I want vision, I go to Dee. When I want something done, I go to Chuck." Russell agreed with this quote and responded: "I think I played a great buffer role." (Russell interview).

[91] Russell interview.

[92] A number of sources remarked that the Hock/Russell team was crucial to Visa's success.

this fee, resulting in large losses for the issuer. Establishing a fixed and fair fee was crucial to the viability of the system, as issuers needed a high-enough fee to make a profit, yet acquirers needed a low-enough fee to be able to offer competitive merchant discounts.

The interchange reimbursement fee has always been a source of controversy in the Visa system. At the core of the issue is a basic question: is it a form of anti-competitive price-fixing, or a necessary aspect of a cooperative payment system? Because this fee is agreed upon by a group of competitors for their benefit, and because it effectively establishes a minimum for the merchant discount fee, many merchants have argued that this is indeed anti-competitive, and on occasion the US Department of Justice has agreed. Visa has typically responded that all cooperative payment systems require these kinds of fees, not only to help cover the participants' operating costs, but also to provide the correct economic incentives for both issuers and acquirers. From Visa's perspective, the fee is a method of obtaining an *economic balance*, providing both acquirers and issuers with sufficient profit while remaining competitive with other payment card systems.[93]

Over the years, this fee has also created controversy between the member banks, as each has tried to tip that balance in their own favor. On average, this fee constitutes ten percent of an issuer's revenue, so banks that specialize in issuing often want to increase this fee.[94] Not surprisingly, those that specialize in acquiring fight to reduce it. When NBI formed, most member banks performed both roles, but as the business matured, banks began to specialize in one role or another, creating a certain partisanship regarding the direction, and eventual segmentation, of the interchange fee.

From a sociological point of view, the controversy surrounding the interchange reimbursement fee is a classic example of a social negotiation process. In any payment system, the participants must answer the basic questions of who pays, who benefits and who gets to decide?[95] There are no "natural" or perfect answers to those questions. The case of cheques presented in the previous chapter also shows how the answers can change over time. In any payment system, the participants must negotiate, taking into account various economic, legal and political factors, in order to arrive at an often temporary solution that is seen as equitable-enough to sustain the system.

The primary issue with the interchange fee under the licensing system was that it was vague and unrealistic to audit, so NBI's officers realized that they needed to establish a fixed

[93]This argument is articulated in the language of multi-sided platform economics in Evans and Schmalensee. For a very detailed theoretical model of four-party payment systems and the need for an interchange fee, see William F Baxter, 'Bank Interchange of Transactional Paper: Legal and Economic Perspectives', *Journal of Law and Economics* 26:3 (October 1983). Note that Visa's change of focus from cost-reimbursement to economic balance is not accidental. As the systems were computerized, the processing costs and risks of extending credit diminished, leading merchants to argue that the fees should also be reduced.

[94]Schmidt interview.

[95]For a discussion of this in the context of the UK's planned point-of-sale electronic funds transfer system, see Howells and Hine. For an excellent analysis of the way different value systems can influence answers to these questions, as well as design priorities, see Rob Kling, 'Value Conflicts and Social Choice in Electronic Funds Transfer System Developments', *Communications of the ACM* 21:8 (August 1978).

interchange fee that was consistent and unambiguous. However, they were faced with a difficult problem—how should they calculate the rate? Setting a proper rate depended on knowing not only the true costs of interchange transactions, but also the entire economics of the system, neither of which were well understood in 1971. The industry was still very young, and most bankcard programs did not delineate their costs by function, much less by transaction type. Many of the small rural banks barely had accounting systems at all.[96]

It was clear that more information was needed to understand the economics of the business. To begin gathering it, NBI established a requirement in late 1970 that all member banks must submit a "certificate of sales" each quarter, which reported their basic operational information such as number of accounts, sales volume, delinquencies, charge-offs, and the like. This information not only helped NBI calculate member fees, but also provided them with the data they needed to build a basic economic model of the business. This information was eventually compiled and summarized by Ron Schmidt into a quarterly "profit analysis report" that was then made available to all the membership starting in third quarter of 1971.[97] Thus, for the first time each member could now see how their program compared to others in the system, as well as how the system was performing as a whole.[98] Furthermore, Hock could use these reports at the Board meetings to prod directors of under-performing banks to improve their programs. As with any network industry, improvements in one element of the system often brings about a benefit to all.

The quarterly reports also served another important function at the time. Throughout 1970 and 1971, industry papers such as the *American Banker* were publishing articles about the huge losses incurred by bankcard programs, questioning whether this endeavor could ever be profitable.[99] According to these articles, many banks were considering giving up on the card business altogether. The quarterly reports provided a concrete rebuttal, showing the member banks that some of them were indeed beginning to turn a profit.[100] Furthermore, each bank could now see how their programs compared to others, creating an incentive for improvement. But most importantly, the reports encouraged the member banks to ask NBI to take an *active* role in helping the individual programs achieve profitability; if a bank was struggling, they could call NBI and a team would visit them to study their procedures and costs, and to share

[96] Schmidt interview.

[97] Schmidt interview. See also Phillip Brooke, 'National BankAmericard Requires Banks' Data for Analysis of Card Profits, Losses', *American Banker*, 6 August 1971, p. 1.

[98] These reports were actually quite detailed. Banks were grouped by the sizes of their portfolios, whether they did acquiring, issuing or both, and whether they used a third-party processor or not.

[99] For example, see the series of articles that begins with Phillip Brooke, 'Banks Reappraise Cards as Losses Mount', *American Banker*, 18 May 1971, p. 1. In the 14 June edition (p. 5), there are a few letters to the editor in response to this series. In one, a bankcard manager wrote "My very candid opinion is that after having operated a credit card operation, I can see nothing in store for the future of this operation but disappointment. Evidently, many banks are reluctant to admit their mistakes and prefer to continue to lose money for their institutions rather than admit they were wrong."

[100] According to the American Banker, one third of the member banks were profitable by 1971. See Kenneth McKenna, 'Banks Find Credit Cards Are Useful...and Costly', *American Banker*, 19 June 1971, p. 3.

best-practices learned from other banks in the system. NBI was thus helping banks *learn from one another* for the benefit of the entire system.

The quarterly reporting data helped NBI understand the general economics of the system, but it still did not contain enough detail about costs to determine an optimal interchange fee rate. Thus, NBI began a detailed research project that would eventually produce what was known as the "Functional Cost Study," first published in the fourth quarter of 1971. While, the quarterly profit analysis reports provided a high-level overview, the functional cost study was a very detailed examination of the functions performed, and costs incurred, by a typical program. Most banks had never examined their programs this closely, so a team comprised of NBI staff and Arthur Andersen consultants visited the bank, observed their operations, and studied their accounting records.[101] This process required about two weeks of research in each bank, plus two or three weeks of additional analysis, but the results provided the bank and NBI with detailed information about costs and problem areas. The bank then knew where to target their efforts, and NBI could construct a more accurate economic model of the system.

However, it should be noted here that the first fixed interchange reimbursement fee was established before these studies and reports, and the model they would inform, could be entirely completed.[102] At the time, NBI was more akin to a high-technology startup and the business was moving so quickly that they could not afford to wait for the results of detailed studies. Schmidt was still in the field researching the member banks when he was asked to estimate a rate. The model he and his Arthur Andersen consultants were building calculated a rate around 2.6 percent, but this seemed too high to Hock. At this time acquirers were sending issuers about 2 percent on the average, so the new rate needed to be closer to this. Hock proposed a rate of 1.95 percent instead. Interestingly, Schmidt remarked that in his opinion, 1.95 percent was suggested instead of an even 2 because it would appear as if it had been arrived at through sophisticated calculation, and not simply decided upon through instinct.[103]

4.4.4 National Advertising

The final step I will discuss in this section is NBI's initial ad campaign, "Think of it as money," which began in April of 1971.[104] From the outside, this may have seemed like any other series of BankAmericard ads, but for Hock, this campaign was his chance to communicate his ideas about the nature of money directly to the cardholders, and he took it very seriously. As Tom Cleveland noted, "There was never a doubt about what was priority one in Dee's mind regarding the relative importance of issues at Visa—how Visa appeared in the market place, from media ads to brochures to corporate business cards, nothing went anywhere without Dee's expressed

[101] These were called "Profit Improvement Teams (PITs)" (Honey interview).
[102] Of course, NBI continued to evaluate and adjust this rate as the model developed.
[103] Schmidt interview.
[104] Chutkow seems to indicate that this tag line was inherited from the BofA, but that NBI developed it further.

approval."[105] NBI reportedly spent $2 million of its own money, plus "substantial additional funds" from NBI member banks to promote the message across a variety of media.[106]

The message of the campaign was twofold. First, it proclaimed to consumers that the BankAmericard is a "modern sensible medium of exchange." The phrase "medium of exchange" is important here—for Hock, the BankAmericard was not just a credit card, it was a *medium of exchange*, and thus a *new kind of money*. This was graphically reinforced by showing a BankAmericard held by a money clip, decorated with a coin. Already we can see Hock pushing the card away from the limited notion of credit, towards the broader concept of value exchange.

The second message was to encourage "responsible use of the card" with the followup tag line, "If we didn't think you could handle it, we wouldn't have given it to you."[107] This was not only an attempt to educate consumers about how to use the card, it was also an attempt at a rebuttal against the various accusations the press and legislators were hurling at the banks throughout the late 1960s. For example, Life magazine ran a scathing cover story against credit cards in 1970, claiming that "In a rush to 'get their plastic on the air,' banks randomly fired off credit cards." It continued:

> They did so, in many cases, with a kind of eager innocence which none of them would have countenanced for a moment in firms with which they did business; a few of them, caught up in the excitement of the unfamiliar chase, seem to have become as blithely careless of consequences as a drunken sailor shooting craps in a Mexican whorehouse on New Year's Eve.[108]

The article went on to quote from a study of 84 card-issuing banks conducted in 1969 for the Charge Account Bankers Association, in which they found "only a fifth of the banks investigated checked the credit background of those to whom cards were mailed," and "a fifth of the banks [in the study] did not even bother to find out if [an applicant] had a job."

Articles like these were naturally shaking consumer confidence in the cards, as well as the banks that issued them, and these ads sought to reassure cardholders that banks were issuing them responsibly. However, it should be noted that they are also an excellent expression of one of Hock's most basic beliefs: the cards offer a valuable service, and consumers are entirely capable of using them responsibly. In a 1979 interview, he stated these beliefs quite clearly:

> Why do people always think of consumers as a mass of ignorant people needing help? If you talk to them as individuals, you will have trouble finding helpless, confused, ignorant people. The average man on the street is not stupid. He is the soundest, most prudent, solvent and dependable part of the economy. Each

[105] Cleveland, p. 26.

[106] 'BankAmericard Starting Ad Campaign Urging Public to Think of Card as Money', *American Banker*, 10 March 1971, p. 1. NBI also received some extra free advertising when Hank Aaron hit his 755th home run directly over the BankAmericard billboard, a scene that was often repeated on television (Cleveland interview).

[107] Ibid.

[108] O'Neil.

> individual is a far better judge of his needs, desires, and financial affairs than many give him credit for.[109]

4.5 Conclusion

In this chapter, we reviewed the core problems with the BankAmericard licensing program in order to show why a new kind of organization was necessary for the system to survive. That organization, which was formed in 1970 as National BankAmericard Incorporated, was designed by Dee Hock, who used it to explore his somewhat unorthodox ideas about the nature of organizations and money. Finally, I discussed some of the first steps taken by that new organization not only to solve some of the pressing problems, but also to generate an understanding of the system as a whole.

In the next chapter we will turn our attention to how NBI and others solved the first of the critical operational problems introduced in this chapter: authorization.

[109] Streeter, p. 75.

Chapter 5

Automating Authorizations: BASE

In the previous chapter, we reviewed the various problems facing the BankAmericard payment system in the late 1960s, showing how the creation of a new, independent organization was necessary for its survival. That new organization, National BankAmericard Incorporated (NBI), provided a structure in which the multiple competing organizations could cooperate to provide a service none could have realistically provided alone. However, the organizational issues were only half of the problem—the operational issues surrounding the authorization, clearing and settlement of interchange transactions were still waiting to be solved.

In this chapter, we will focus on how NBI and others addressed what was considered to be the most critical of these operational problems: authorization. By 1970, The slow and cumbersome manual authorization process had become what Hughes called a "reverse salient."[1] I will begin by briefly reviewing that concept, and describing the critical problems, along with the implied solutions, identified by the industry leaders. I will then discuss how various organizations developed solutions for these critical problems, ending with the creation of NBI's national electronic authorization system, BankAmericard Authorization System Experimental (BASE), which was first put into operation in April of 1973.

5.1 Authorization as a "Reverse Salient"

Of the many operational problems facing the bankcard systems in general, and the BankAmericard system in particular, authorization was seen as the most important. This was for two reasons: first, the delay and hassle involved in obtaining an authorization was beginning to affect consumers' desires to use the card, and merchants' willingness to accept it; second, the floor limits, which were intended to ease the former concerns, were simultaneously making it difficult to control the increasing levels of fraud, which was in turn further eroding consumer confidence. Without the continued participation of both cardholders and merchants, all the

[1] Hughes, *Networks of Power*, p. 79.

bankcard systems would quickly collapse. This is how one of the industry leaders described the situation:

> Vital—critical—all important. These words you have heard to describe the urgency that surrounds our troubles with authorization. And I suppose what frustrates so many of us is the conviction that this one trouble can be resolved. If authorization is so vital, so critical, if it can generate more business for the card, if it can lead to a reduction in fraud and credit losses and effect greater economies in our operation—then why, as the little boy says, can't I have one of those things, Daddy?[2]

In Hughes's terms, the authorization process had become a "reverse salient," that is, an element holding back, or even thwarting, the development and growth of the overall system.[3] Hughes developed this concept while trying to understand how complex sociotechnical systems tend to evolve. He observed that in many cases, one component will "fall behind or out of line" with respect to others as the system progresses towards a specific, intermediate goal. Borrowing a term from military history, he called this phenomenon a "reverse salient," which he thought better captured its contingent and chaotic nature than "bottlneck" or "disequillibrium." Hughes noted that when a reverse salient occurs, numerous actors will focus their attentions upon this component, and take "remedial action" to correct it.

However, it is important to remember that a reverse salient does not necessarily *prohibit* the operation of the current system, nor some small amount of potential growth. By arguing that authorization had become a reverse salient, I do not mean to imply that the bankcard systems were incapable of operating without immediate "remedial action." Throughout the early 1970s, both the BankAmericard and Interbank systems continued to function, albeit badly, with the existing manual authorization process. However, the important point is that these systems could not have achieved any significant growth, and therefore would have eventually fizzled out or fallen to a competitor like American Express, without some sort of remedial action.

Hughes argued that when confronted with a reverse salient, actors seek to correct it by constructing one or more "critical problems," the articulation of which often guides them towards, or even directly implies, certain solutions.[4] In the case of bankcard authorization, the industry leaders constructed two critical problems. First, the local *authorization decision* was too slow, and often not available, because it required *human intervention*. Second, *interchange*

[2] Phillip Brooke, 'Quigley Describes Operation of Omniswitch; Urges Cooperation in National Development', *American Banker* (11 August 1971), p. 6.

[3] Hughes, *Networks of Power*, p. 79.

[4] MacKenzie further points out that historians should never treat reverse salients or critical problems as "given, independent of the actors involved." In the midst of events, actors may not agree on what goal should be obtained, what is hampering progress toward it, and how it might be solved. Causation may also run the other way—actors may formulate critical problems based on what they can accomplish, and then argue for a corresponding reverse salient. See MacKenzie, 'Missile Accuracy', pp. 197–199. In the case of authorization, there seems to have been rather widespread agreement that it was a significant problem area, though I will note the few minor disagreements regarding how it should be solved.

authorizations were too slow because they required manual "two-legged" calls or telexes between the acquiring and issuing centers. The solution implied by the first problem was replacing the human authorizers with automated, computerized logic. The solution implied by the second problem was enabling electronic communication between the centers, either computer-to-computer, or computer-to-terminal. Building an effective nationwide card system, and its accompanying authorization system, ultimately required solving both problems.[5] However, local authorizations were automated first, primarily because they did not require cooperation and coordination between independent, competing organizations.[6] Therefore, we will begin by charting the various ways in which organizations automated their local authorization decisions.

5.2 Automating Local Authorizations

National Data Corporation (NDC) of Atlanta seems to be the first processor to offer some sort of computerized authorization to its subscribers.[7] NDC began operations in 1968 as the primary processor for the Chicago-area Interbank members, but eventually expanded to handle processing for a number of merchant-specific card systems, oil industry cards, and two NBI member banks. NDC also provided "after-hours" authorization services for banks that wanted to operate their own center during the normal working hours; after the bank's center closed, they would switch their phone network so that calls would go to NDC instead. Because NDC was a processor for many issuers and acquirers, they took the approach of centralizing all cardholder data onto one computer system from which they could make automated authorization decisions. Merchants called one of NDC's four regional authorization centers, where less-skilled, clerical operators keyed the transaction information into terminals, which in turn communicated over 2400 baud modems with NDC's central computer (a UNIVAC 494) in Atlanta. The key aspect of this system was that the authorization decision was entirely automated; subscribers could define the rules used to make the decision, but the rules were then executed without human intervention. Most subscribers opted for *negative authorization*, meaning that card numbers were simply checked against a "derog" file, which contained accounts on which purchases should not be authorized. However, some took advantage of NDC's scheme for *positive authorization*, which maintained an "open to buy" amount for the particular account, as well as a history of transactions over the previous seven days. The automated decision significantly sped up the authorization process, resulting in an average local authorization time of just twenty-two seconds. However, NDC's system could authorize transactions only for cards issued by one of their subscribers—interchange authorizations required an additional, "two-legged" call to the

[5] It should be noted that some argued that the second problem could be made moot by consolidating all issuers under one, centralized processor. However, this was not politically feasible, as the issuers feared this would restrict their ability to compete with their rival banks.

[6] Recall that in a local authorization, the acquirer and issuer are the same organization, or are served by the same processor, so political issues are largely removed.

[7] 'National Authorization Joint Feasibility Study Final Report' (29 January 1971), p. 8.

issuer's center.[8]

Credit Systems Incorporated (CSI), the regional processor for Interbank members in Missouri, Kansas, Illinois, and Kentucky, also began offering computerized authorization to its subscribers in early 1971.[9] Their system was similar in concept to NDC's, but was augmented to offer a completely automated interface for the merchant. Merchants could use a touch-tone telephone, a relatively new device at the time, to enter the transactions details, guided by voice prompts. The computer system (an IBM 360 with a tone frequency decoder) translated the tones into an authorization request on their local cardholder data and used the response to select the appropriate pre-recorded message that was played back to the merchant. Interestingly, the system itself never responded with a pre-recorded denial; if the automated rules could not authorize the transaction, the call was switched to a human authorizer who reviewed the transaction and account details on a terminal and made a final determination.[10] Similar to NDC, CSI's system could authorize transactions only for cards issued by one of their subscribers, but in the case of an interchange authorization, the merchant phone call was automatically switched to the issuer's center, and the merchant spoke directly with the issuing authorizer.

In June of 1971, Omniswitch began what seems to be the first test of merchant point-of-sale (POS) terminals with a bankcard program.[11] Like the charga-plate system mentioned in Chapter 3, these counter-top electronic devices helped to reduce data entry errors by directly reading and transmitting the card information.[12] The ABA had publicly endorsed the magnetic stripe (often abbreviated "magstripe") in early 1971 as the preferred method for making the cards machine-readable and defined a format for encoding the account information upon it.[13] As a result, a number of manufacturers began producing terminals in 1971 designed to read the ABA format and transmit the card data across standard telephone lines.[14] However, these early

[8] 'National Authorization Joint Feasibility Study Final Report', pp. 8–9.

[9] 'Fast Credit Card Authorization is Offered to Banks and Merchants by CSI', *American Banker* (6 January 1971), p. 8.

[10] Unfortunately, the article did not mention a reason for this design choice, but there are two possibilities: first, because a denial embarrasses and erodes the confidence of a legitimate cardholder, such a decision should be verified by a human; second, it may have been culturally inappropriate for a denial to come from an automated, impersonal recording.

[11] C Frederic Wiegold, 'Omniswitch Tests System of Merchant-to-Bank Authorization to Aid Card Use, Reduce Fraud', *American Banker* (18 June 1971), p. 1. See also, 'Omniswitch Tests Merchant-to-Bank Authorization System', *Payment Systems Newsletter* (July 1971), p. 4. Whether Omniswitch was actually the first to test POS terminals is somewhat unclear. The news articles imply this but are also vague enough to create doubt. I have not yet found an earlier report about POS terminals in the banking news sources. Tom Schramm, Omniswitch's VP of operations, noted that he thought they were at least one of the first to test them, if not the first (Schramm interview).

[12] Note that these terminals performed authorization only. Merchants were still required to use a standard imprinter to complete a paper sales draft for clearing and settlement.

[13] 'Magnetic Stripe for Credit Cards Urged by ABA Unit', *American Banker* (16 February 1971), p. 1. See also 'ABA Adds to Guides For Magnetic Card Coding', *American Banker* (18 March 1971), p. 1. Interestingly, the airlines favored putting the stripe on the front of the card, but the ABA decided that this would "would seriously deface the logo of the bank plans," and chose instead to put it on the back. However, the Japanese banks put the stripe on the front of their cards for many years (Conway interview).

[14] Manufacturers included IBM, Data Source, Addressograph, and Transaction Technology Incorporated, a subsidiary of First National City Corporation (now known as CitiCorp).

terminals were quite bulky and expensive, ranging from $500 to $1,000 each, so they were appropriate only for larger merchants who generated numerous card transactions. Omniswitch's system will be discussed in more detail in the next section, but it should be noted here that the POS terminals further decreased the average authorization time to just fifteen seconds.[15]

NDC also added POS terminals to their authorization system in the summer of 1971, but instead of reading a magstripe on the back of the card, their terminals optically scanned the embossed characters on the front. These terminals, manufactured by Data Source Corporation, were adopted mostly by the oil companies, who felt little need to follow the ABA magstripe standard. With these terminals, NDC's average authorization time was reduced to just seven seconds.[16]

In October of 1971, City National Bank and Trust (CNBT) of Columbus, Ohio began what they called an "electronic funds transfer pilot test."[17] This test used similar technology to those already mentioned, but as the name suggests, it had a much more ambitious objective. It was not merely a test of computerized authorization equipment; it was a test of the very idea of full electronic funds transfer (EFT). A spokesman described their intent:

> What we really hope to do is to peek into the future and learn the sequence of social and technological developments that will bring about a society where most sales involve the electronic transfer of data and funds, instead of cash and checks.[18]

CNBT was an NBI member bank, and one of the few at that time that shared Hock's vision for an electronic value-exchange system. However, it was not entirely clear in 1971 if the public would accept such an arrangement. In fact, the Federal Reserve Bank in Atlanta commissioned a study on the public reaction to electronic payments earlier that year, and the final report stated that the public was overwhelmingly against any kind of electronic payments system.[19] However, the head of the research group surmised that this had more to do with a lack of understanding as to how such a system would actually work. The lesson was that any system developed would require extensive consumer education in order to gain acceptance. A key element of CNBT's test was to determine if a particular set of consumers would accept the idea of electronic payments and use them to the exclusion of cash and cheques.

The test was fairly small, and was conducted in the favorable location of Upper Arlington, a wealthy suburb of Columbus. CNBT issued two different cards to 20,000 residents and

[15] Interestingly, Omniswitch found that this response time was almost too fast for merchants, and built in a "delay factor... to facilitate human reaction time." See 'Omniswitch Tests Merchant-to-Bank Authorization System', p. 5.

[16] 'NDC Credit Authorization Pilot Underway', *Payment Systems Newsletter* (July 1971), p. 7.

[17] 'Electronic Funds Transfer Test Announced by City NB&T', *Payment Systems Newsletter* (July 1971), p. 5. See also Phillip Brooke, 'City NB&T, Columbus, to Test Point-of-Sale Card Authorization, Data Capture in Suburbs', *American Banker* (14 July 1971), p. 1. CNBT is now known as BancOne. They remained at the forefront of technological innovation, providing the bank processing side of the Merrill Lynch Cash Management Account (CMA), the first investment account that could be accessed using cheques or a payment card. See Nocera, pp. 159–199.

[18] C Gordon Jelliffe, president of CNBT, quoted in the *Payment Systems Newsletter* (July 1971), p. 6.

[19] Phillip Brooke, 'Public Opposes Shift to Electronic Payments', *American Banker* (9 July 1971), p. 1.

installed 58 IBM POS terminals at 29 merchant locations. The first card was a modified form of their standard BankAmericard, featuring a magstripe on the back. With it, cardholders could make credit purchases at participating merchants, where all purchases were electronically authorized using automated logic regardless of amount (a condition known as a *zero floor limit environment*). Although merchants still completed a paper sales draft, it was reported that the authorization request actually caused an immediate debit to the cardholder's credit line, and a corresponding credit to the merchant's account.[20] Thus the paper sales draft functioned mostly as a receipt for the customer, and not a claim on funds, as the transaction was already cleared and settled electronically.[21]

The second card was specially designed for use in the new Docutel "Automated Total Teller," which could accept deposits/payments and transfer funds between accounts, in addition to dispensing cash. With the combination of both cards, consumers could significantly reduce their use of cheques, limit the amount of cash carried at any one time, and perform nearly all their banking tasks without ever visiting a physical branch.

After four months, the bank declared the test to be a success, both technically and culturally, but the degree of its success is debatable. A spokesman reported that the "bank is finding not only that the new technology involved works, but also that the public will *adapt to change*."[22] The language here was no doubt chosen carefully—the public may be willing to "adapt" to change, but very few were actually demanding it. Although the merchants were enthusiastic, the cardholder reaction was decidedly lukewarm—only 20,000 transactions were processed within the first 100 days, just one for each card issued. It was clear that consumers were still using cash and cheques for most purchases.

Despite CNBT's inability to convince their customers of the benefits of full EFT, these early systems did convince issuers of the feasibility and benefits of replacing human authorizers with computerized logic. Many organizations continued the trend by automating their local transactions in a similar manner throughout the rest of the 1970s. However, interchange authorizations still required a manual, two-legged call or telex, resulting in multiple-minute authorization times for those traveling outside of their area, and limited non-local fraud detection for the issuers. One system that promised an innovative solution to this problem was Omniswitch.

5.3 Automating Interchange Authorizations

The Omniswitch organization was originally formed in 1969 to provide New York merchants a single, centralized authorization service for all Master Charge cards. The New York metropoli-

[20] *Payment Systems Newsletter*, July 1971, p. 5. See also Phillip Brooke, 'Bank, Shopper, Stores Like Electronic Funds Transfer in Ohio Test', *American Banker* (28 January 1972), p. 1.

[21] Whether merchants were still required to deposit the paper drafts is not mentioned.

[22] Phillip Brooke, 'Bank, Shopper, Stores Like Electronic Funds Transfer in Ohio Test', *American Banker* (28 January 1972), p. 1. Emphasis added.

tan area was one of those where banking regulations and competition made it relatively common for the merchant and cardholder to be represented by different banks. Most of the Interbank members in that region belonged to the Eastern States Bankcard Association (ESBA), who also performed their processing, but the First National City Bank (FNCB, later Citibank) did not, as they had converted to Master Charge from their proprietary "Everything Card" and thus already had their own processing center. Omniswitch eliminated the need for merchants to determine which center to call, providing them with one point of contact for all Master Charge authorizations, regardless of issuer. Omniswitch began offering this service in June of 1970.[23]

Because FNCB was uninterested in transferring their cardholder data to ESBA's computer, Omniswitch developed the innovative approach of "switching" the authorization request messages to the appropriate computer system, similar to how a networking switch routes packets to the appropriate node on a computer network.[24] Merchants called Omniswitch's data center in Lake Success, NY, where operators keyed the requests into an interactive computer terminal. The Omniswitch computer (an IBM 360/40) then used the first few digits of the card number to determine the appropriate destination: the ESBA computer, which was literally across the room; or the FNCB authorization computer, which was about 9 miles away.[25] The Omniswitch computer transmitted the messages to the issuer's computer system, where they were processed without human intervention. One of three replies could be returned: an approval with a corresponding code; a denial; or a "referral," meaning that the operator should connect the merchant to the issuer's center to obtain further details. As noted earlier, Omniswitch began testing POS terminals in 1971, and also offered a voice-prompted touch-tone interface.

This switching technique had two important advantages over the consolidation approach taken by organizations such as NDC. First, it made it politically easier to expand the geographic reach of the system, as banks and processors could maintain control over their cardholder data and local authorization decisions, but utilize Omniswitch to process interchange authorizations at nearly the same speed.[26] Second, it also made it technically easier to expand the system, as supporting a new processor required only the development of a relatively simple bridging program, what we would today call a "driver" or "provider." Processors could and did use a wide variety of computer hardware and software for their local processing, but they could still communicate through Omniswitch to other issuers, as Omniswitch provided all the necessary

[23] Information on Omniswitch was collected through interviews with Tom Schramm, former VP of Operations. Schramm also provided reports, presentations, and news clippings from the time, allowing me to verify his comments with the written record. The most detailed news story on Omniswitch is Phillip Brooke, 'Quigley Describes Operation of Omniswitch; Urges Cooperation in National Development', *American Banker* (11 August 1971). Schramm came to NBI in 1974.

[24] To be clear, Omniswitch's innovation was the switching of *authorization* request messages, not the switching of digital messages in general. For the development of packet switching, see Janet Abbate, *Inventing the Internet* (Cambridge, MA: MIT Press, 1999).

[25] Schramm interview.

[26] The decentralization of cardholder data also pleased a US Senate sub-committee that was concerned about potential loss of privacy due to data consolidation. See 'Omniswtich: A Cornerstone for National Authorization Through Interbank' (14 December 1971), p. 20.

protocol conversion.

Although Omniswitch began with only two nodes, it quickly expanded to five by the end of 1971. Because most of these nodes were themselves multi-bank processors, the Omniswitch network effectively provided interchange authorization services for 286 banks and 100,000 merchants spread across 11 states. Throughout 1971, Omniswitch routed 4.5 million transactions, and handled a peak of 2,500 merchant calls per hour at their Lake Success authorization center.[27]

Omniswitch was able to have this multiplicative effect because of the specific history of Interbank's formation and growth, which was quite different from that of the BankAmericard system. The Interbank system began as an alliance between regional processors, each of which served many organizations, while the BankAmericard system began by licensing individual banks. From an organizational perspective, Interbank was a network of networks, while BankAmericard was a simpler star network of individual nodes.[28] Thus, Omniswitch could provide interchange authorization for numerous banks by connecting only a few processors. From a technical perspective, this reduced their complexity, as well as their transaction load. Most of the authorizations could be handled within the local processor, and only a small percentage of those required interchange.[29]

Omniswitch continued to expand, and in January of 1972 they agreed to add a new node that dramatically increased their scope: NDC, the large processor mentioned earlier.[30] With the NDC connection, Omniswitch now provided interchange authorization for 2,800 banks, which comprised 72 percent of the entire Interbank system. Although Omniswitch and NDC formed a new organization called Nataswitch to govern the connection, it was technologically just another node on the Omniswitch network.

Omniswitch was actually designed from the beginning to be a dependable nationwide authorization system, and thus developed many of the features that would later appear in the systems built by NBI and Interbank. For example, although issuers typically performed positive authorization, Omniswitch also maintained a file on its computer that could be used for negative authorization if the communication lines failed, or if the issuer's system was unavailable or unresponsive. Any authorizations made under these conditions were logged and sent to the issuer electronically when their system became available again. Omniswitch also maintained a redundant IBM 360 as a hot backup and could switch to it almost instantaneously if the primary system failed.[31]

[27] 'Omniswitch: A Cornerstone', pp. 7–8. See also Daniel Quigley, 'Omniswitch System Described as Reality, Not Concept', *American Banker* (1 December 1971), p. 8A.

[28] Although a few NBI members used third-party processors (such as NDC) in the early 1970s, the vast majority did not.

[29] There were of course localized exceptions to this, as in the New York area. Also, a few acquirers contracted Omniswitch to handle their merchant calls directly via inbound Wide Area Telephone Service (WATS) lines.

[30] Phillip Brooke, 'Omniswitch, National Data Joining System', *American Banker* (4 January 1972), p. 1.

[31] Schramm interview. Confirmed in Brooke, 'Quigley Describes Omniswitch'.

5.4 The Joint National Authorization System

Omniswitch established a model for how to build a decentralized national authorization system by networking together the existing local centers and switching authorization requests between them. This was something NBI, Interbank and American Express all wanted to provide, and during Omniswitch's first year of operation, their representatives met to discuss the possibility of developing one joint computerized system that could provide authorizations anywhere, at any time, for any card program.[32] From a systems engineering perspective, it seemed wasteful to develop separate authorization systems for each card network, as they would all need to do essentially the same task. It seemed far more "logical" to combine efforts and build one shared system.

The participants formed an advisory team, which investigated the technical feasibility of such a system and compiled their final report in January of 1971.[33] In it they concluded that such a system was indeed feasible, and recommended using the message-switching technique developed by Omniswitch. They estimated that a single switching center could be operational in eighteen to twenty-four months, and two more could be added by 1976, to handle the expected increase in volume.

A shared system may have been technical feasible, but was it politically so? Hock was vehemently opposed to it because he thought that a single, shared system was antithetical to his vision:

> From my perspective, neither the institutional nor technical thinking made sense. It was just another attempt to centralize power and control.... Creating a single, monopolistic, electronic payment system seemed... [to be] an attempt to wrap the substance of the future in order to perpetuate past forms. It was contrary to all my beliefs about the nature of organizations and the possibilities inherent in electronic communications. Exchanging authorization information and monetary value in the form of electronic particles ought to be a highly decentralized, competitive business. Trying to design and impose a single, monolithic system on such an essential flow of information seemed absurd.[34]

However, it is equally possible that Hock was simply opposed to any such system that he did not completely control. Hock's more ambitious goals required a large-scale computer network similar to what the joint feasibility study was proposing. If he did not control it, he would lack the ability to dictate the development of the system to ensure that his goals could be

[32] This was actually not the first or last time a joint system was proposed. As NBI was forming, Bank of America and American Express proposed developing a joint system to which they would then sell access, but found few takers. The ABA Monetary and Payments System (MAPS) planning committee also proposed a centralized system run by the large commercial banks, but also failed to generate any action. See Hock, *One From Many*, pp. 164–165; and Phillip Brooke, 'MAPS Planners, in Final Report, Urge Payments System Free of Fed Control', *American Banker* (3 May 1971), p. 1.

[33] 'National Authorization Joint Feasibility Study Final Report'.

[34] Hock, *One From Many*, p. 165.

achieved. Once a shared system was in place, it would also become difficult to convince banks to fund a second, private network to deliver the services Hock ultimately wanted to provide.

At the time, NBI's Jack Dillon also claimed that the US Department of Justice was starting to object to the idea of the banks and the T&E companies cooperating to build a joint system.[35] This is not mentioned in Hock's autobiography, nor Visa's corporate history, but it is entirely possible that the DOJ would object to such a system. Because acquirers (and ultimately merchants) pay to authorize transactions, such a jointly-owned system would result in a fee set by a number of competitors. The DOJ tends to interpret these activities as price-fixing and thus a breach of antitrust regulations. Because those who lose an antitrust lawsuit are required to pay treble damages, many organizations will choose to abandon efforts that the DOJ even hints might become questionable.

At the June 1971 Charge Account Bankers Association conference in the Bahamas, Hock announced that NBI would abandon the joint effort, take a "unilateral approach" to national authorization, and develop its own private system.[36] Interbank followed suit, and the joint effort effectively dissolved.[37]

5.5 NBI's BASE

Having decided to develop their own system, NBI did what most firms at that time did: hire a consultant and put out a Request For Proposals (RFP). Of the sixty vendors interested, NBI invited twenty-one to attend a meeting on 8 October 1971, where NBI presented their needs. At this time, NBI was primarily interested in buying an existing system in order to offer a national authorization service in the shortest amount of time possible. In fact, the RFP required that any proposed system must be operational within twelve months.[38]

Although the requested system, which was given the tentative name of "BankAmericard Authorization System Experimental (BASE)," was primarily intended to provide authorization services, the RFP revealed that Hock already had much larger plans. National authorization was merely "the first phase of a more comprehensive nationwide bank information processing system."[39] Building BASE would create a computer network connecting the seventy-six ex-

[35] Phillip Brooke, 'NBI Plans to Sign Pact For Nationwide Authorization System by Year's End', *American Banker* (4 November 1971), p. 1.

[36] 'Card Groups Take Own Authorization Paths', *American Banker* (29 June 1971), p. 1.

[37] Unfortunately for Omniswitch, CSI was given the contract to build the Interbank National Authorization System (INAS). According to Schramm, the reasons were "entirely political." FNCB, who was part-owner of Omniswitch, had also designed their own competing system that they wholly-owned. FNCB represented both Omniswitch and their own system in the selection process, and their conflict of interest caused them to downplay Omniswitch in favor of their system, which was not even operational at the time. CSI eventually won the contract because their system was functioning, and their leader was also leading the overall selection process. Omniswitch was eventually abandoned once INAS became operational.

[38] Phillip Brooke, 'NBI Plans to Sign Pact For Nationwide Authorization System by Year's End', *American Banker* (4 November 1971), p. 1.

[39] Brooke, 'NBI Plans'.

isting BankAmericard centers and processors spread across the nation, and once that network existed, authorization was only one of the many possible services NBI could offer. From the beginning, Hock was intending NBI to be the electronic hub through which all electronic value exchange transactions flowed.

Thirteen vendors submitted proposals in November 1971, and at the time NBI was confident it could offer a nationwide authorization service by the second quarter of 1972.[40] However, in late February 1972, NBI announced that it had rejected all thirteen proposals as none "satisfied enough of [their] functional needs."[41] This very careful statement attempted to gloss over a much worse reality that Hock later admitted: even the best bid was several times more than their allocated budget, twice as long as their desired schedule, and "no vendor was willing to warrant the performance of the system."[42] Hock was told that this was customary in the computer industry and that he should just go back to the NBI Board and ask for more money and time. But that would have been antithetical to Hock's personality, not to mention damaging to his reputation, and quoting Emerson's "Trust thyself," he declared that NBI would design and build their own system within the budget and timeframe already approved by the Board.[43]

5.5.1 Acquiring the Talent

However, NBI needed help, as they did not have the necessary staff in 1972 to design and build a large-scale computer system.[44] They contracted with TRW, one of the thirteen vendors who had submitted a proposal, to design and coordinate development of the system. NBI also hired Aram Tootelian away from TRW, where he had been the General Manager of the Information Systems Division, to lead NBI's new Systems and Technology Division. Tootelian had an almost serendipitous background: while working for IBM, he had learned about the Sabre online reservation system built for the airlines, designed and built several bank information systems involving long-distance terminals, and had worked on the TRW Credit Data account, which provided online credit information to banks. While at TRW, he also learned the high-reliability and fail-safe techniques they had developed for top-secret aerospace projects, and even built a computerized simulation of the US payments system for the Federal Reserve. Tootelian remarked that he was the "only person" who could have built BASE, because his background had provided all the required skills and techniques.[45]

Tootelian then hired two others that would become key players in the design and construction of BASE. Dave Huemer, who had been at California Computing after receiving a doctorate

[40] Brooke, 'NBI Plans'.

[41] Phillip Brooke, 'BankAmericard Schedules April Start for 24-Hour National Card Authorization', *American Banker* (21 August 1972), p. 1.

[42] Hock, *One From Many*, p. 171.

[43] Ibid. Although it was risky, Russell agreed with Hock, provided they hired the right people to build and operate it (Russell interview).

[44] In fact, NBI's complete staff totaled less than 20 employees.

[45] Tootelian interview. Biographic data also in Brooke, 'BankAmericard Schedules April Start', p. 1.

in mathematics from Cal Poly Pomona, came as the Director of Systems Operations, responsible for the data center and all its associated hardware. David Goldsmith, who was conducting sales and executive education for IBM after earning his MBA, joined as the Director of Development, responsible for software development and member bank training.

Hock assigned Tootelian's team a formidable task: establish a nationwide computer telecommunications network; install terminals in each of the BankAmericard centers around the country and train each center's staff on how to use them; obtain all the necessary computer and networking hardware; build a data center to house that hardware; install four regional concentrator minicomputers; write, test and debug software that could provide online switching of authorization messages twenty-four hours a day, seven days a week; and staff a call center that could take merchant calls after the local card centers closed for the night. And all of this had to be done by 1 April 1973, just nine months away.

Despite the usual association with that day of the year, Hock was not joking. Throughout his time at NBI/Visa, Hock never allowed a technology project to extend over twelve months. He maintained that if you give computer people more time, they will just consume it, so he always insisted on shorter projects with uncompromising deadlines. As the overall payment system grew, deadlines became even more important, as changes to the central computer system necessitated parallel changes to the members' computer systems as well. Many sources from the systems area indicated that they felt this adherence to short project cycles was a key element of their success.[46]

5.5.2 Design of the System

At the heart of BASE was a Digital Equipment Corporation (DEC) PDP-11/45 minicomputer, housed in NBI's newly constructed data center in San Mateo, California, just down the peninsula from their San Francisco headquarters.[47] This central computer acted as a real-time switch for authorization requests, which could either come from acquiring authorization centers, or directly from the electronic cash register systems of large national merchants.[48] Acquiring centers entered their request into a Harris-Sanders model 804 terminal (often shortened to "Sanders terminal"), which could queue up to ten messages in its internal memory. Periodically a regional concentrator, which was a DEC PDP-11/20, polled the terminals in its area, collected all pending requests and sent them on to the central switching computer.[49] The switch

[46] For a treatise on scheduling techniques for software projects, see Frederick P Brooks, *The Mythical Man Month: Essays on Software Engineering*, Anniversary edition (1975; reprint, Reading, MA: Addison-Wesley, 1995). Although Brooks encouraged developers to be more aggressive about defending how long it takes to write a program properly, he also acknowledged that shorter schedules with clearly defined milestones greatly increased the chances of staying on schedule and delivering something genuinely useful.

[47] Derman and Tootelian interviews. See also Brooke, 'BankAmericard Schedules April Start', p. 1. Hock had originally requested that the computer be installed in a closet near his office in the BofA building, but was eventually convinced that this would be impractical (Totten interview).

[48] Ibid.

[49] The network was supplied by AT&T and operated at a "blazing" 2400 bits per second.

would determine the issuer from the first few digits of the card number, and forward the request to the appropriate destination.

The issuers also had a Sanders terminal in their authorization center, on which new requests would appear. However, because many issuers experienced such low volumes at this time, it was common that nobody would be sitting in front of the Sanders terminal to see the incoming authorization request. To remedy this situation, NBI asked Sanders to add a small bell to the terminal, the clapper of which could be triggered electronically when a new request arrived. Convincing Sanders to add the bell proved to be one of NBI's more difficult tasks in building BASE, but Sanders eventually relented, and issuing authorizers learned to jump to their terminals when they heard the bell ring. After making the authorization decision (either by using their own computer system or printed reports), the issuing authorizer typed the response into the terminal, which was then routed back to the acquirer's terminal through NBI's central switch.

Although most NBI member banks had not yet automated their local authorizations, a few of the more technically-advanced banks had done so, and these banks wanted a direct CPU interface to BASE.[50] However, connecting a bank's computer system to BASE was problematic at two levels. At the networking level, the DEC PDP-11 used a different communications protocol than the IBM mainframes used by most banks. To enable the systems to communicate, NBI developed a program called the Bisynchronous Communications Module (BCM), which translated the DEC protocol into something the bank's computer could understand. This allowed the computer systems to talk, but what they talked *about* was still an issue. At the application level, the format used by BASE for authorization request and response messages was typically different from those the banks used in their systems. Thus, banks were required to develop a second program that translated BASE's messages into those used by their internal systems. This allowed BASE to treat all end-points on the network in a uniform manner.[51] Over the next few years, NBI helped all the member banks and processors develop such CPU interfaces, as they greatly reduced the overall authorization time.

On the merchant side, a few of the large, national retailers also had direct connections into BASE.[52] From a business perspective, these merchants were still represented by an acquiring bank, to whom they also paid merchant discounts. However, from a data processing perspective, their electronic cash register (ECR) systems were connected directly to BASE, and thus they were effectively their own processor. As we shall see, this dichotomy eventually caused merchants to question why they were paying a fee to a bank that seemingly did little work for

[50] Only four banks had direct interfaces when BASE began operating. See Phillip Brooke, 'NBI Moves Toward Electronic Interchange of Bank Card Sales Drafts to Reduce Costs', *American Banker* (29 August 1973), p. 1.

[51] Fojtik and Derman interviews. Eventually, these message formats were changed and standardized through ANSI and ISO. See Chapter 8.

[52] Brooke, 'BankAmericard Schedules April Start', p. 1. The article makes no mention of who was connected in this way, just that some were. However, it should be noted that these were not the major department stores, as they did not accept bankcards at this time.

them. This climaxed when NBI (by then renamed to Visa USA, Inc.) signed up JC Penney, then the third largest department store chain in the US, in exchange for a direct business and data processing relationship with Visa, bypassing the merchant acquiring bank altogether.[53]

BASE was also built from the beginning to offer a constant, reliable service. Nearly every piece of the system had a redundant backup, including standard dial-up telecommunication lines that could be used if the leased lines failed. However, even if the central computers and telecommunications were functioning, there was no guarantee that the issuer would respond to the request in an acceptable amount of time. To ensure a reliable service, BASE also had the capability to "stand-in" and approve transactions when the issuer did not respond in a timely manner. The central computer could use a negative file, as well as a file containing recent account activity, to approve or deny requests according to rules established by the issuer. The issuer was then notified of all activity when it's system came back online. Interestingly, there was also a "VIP file" containing numbers that should automatically be approved without consulting the issuer. Both the negative and VIP files were maintained through the Sanders terminals, or via tapes mailed to NBI's data center.[54]

Additionally, NBI also had the ability to stand-in for acquirers after-hours, or during especially busy periods. When an acquirer's center closed for the night, they could switch their phone network so that authorization calls went directly to NBI's data center in San Mateo. NBI's own authorization operators handled the calls, entering the requests into their own Sanders terminals. Some smaller centers found this so convenient that they simply contracted with NBI to handle all their authorization calls.[55]

It is also interesting to note how the distribution of work in BASE reflected Hock's organizational philosophy. The local card centers maintained their autonomy, and NBI provided only those services that local centers could not accomplish on their own. Each BankAmericard center or processor was responsible for interacting with their merchants, and maintaining their cardholder accounts. All local authorizations could still be handled by the center itself, but interchange authorizations could be switched through BASE. This design was not only politically attractive, it was also very efficient—the transaction load was distributed amongst the centers, and only the interchange traffic came through BASE, which was still a small percentage of transactions at this time.

NBI's decision to use the DEC PDP-11 instead of an IBM mainframe was somewhat surprising and deserves some comment. Most of the banking industry had standardized on IBM hardware, but Hock was in the middle of a personal feud with "Big Blue." Sources differ as to the exact cause, but all agree that IBM promised something to the fledgling NBI but later

[53] Streeter. The details of the JC Penney deal will be discussed in Chapter 9.

[54] Tootelian and Derman interviews. See also, Brooke, 'BankAmericard Schedules April Start', p. 1.

[55] Not surprisingly, this created some tension between NBI and the various third-party processors. Organizations such as NDC already offered after-hour services to card centers, and they considered NBI to be encroaching on their business. NBI eventually discontinued this service when processors like NDC and First Data Resources offered to handle calls for both Master Charge and BankAmericard authorizations, something NBI could not do.

reneged, which infuriated Hock.[56] At the time NBI was a relatively small, unimportant account for IBM, but in Hock's mind, he was the future of banking and global value exchange. Although Hock vowed that he would never do business with IBM again, as we shall see, he changed his mind a few years later when the transaction volume proved too much for even the most powerful of the PDP-11 series.[57]

One cannot help but notice the similarities between the functional design of BASE and that of Omniswitch. However, it does seem that this was a case of independent, parallel invention, rather than one of direct influence.[58] Although Tootelian was aware of Omniswitch, and although NBI's management would have certainly learned the essential details of Omniswitch from the Joint Authorization System Feasibility Study, Tootelian maintains that the design of BASE came entirely from his past projects and what he considered to be "common sense."[59] In many ways the similarities are not surprising—switching authorization requests would have been the only politically acceptable solution for the BankAmericard system, and the ability to stand-in for issuers is a feature that naturally follows from the desire to maintain a constant, reliable service. Although it may have been technically feasible to centralize cardholder data into one computer system, as NDC had done, NBI member banks would have never agreed to that loss of control.[60]

It should also be noted that there was a significant difference in network scale between BASE and Omniswitch. As discussed earlier, Omniswitch could provide interchange authorization switching for the majority of the Interbank members by connecting only a few large processors. BASE on the other hand, initially needed to connect seventy-six individual bankcard centers and processors, and be able to accommodate more as the membership continued to increase. More members also meant a higher percentage of interchange transactions, and the overall system volume was growing at an alarming rate. During the construction of BASE, the sales volume generated in the US BankAmericard system grew by thirty-two percent, and the first quarter of 1973 set a record with a staggering forty-one percent growth from the same quarter the previous year.[61] This quickly inflating scale posed unique challenges that required different approaches, such as the regional concentrators, special queueing logic, and

[56] For Hock's recollection see, Hock, *One From Many*, pp. 172–173.

[57] See Chapter 7. Tootelian also remarked that DEC could deliver the hardware when they needed it, while IBM could not, and that the DEC was a simpler, yet perfectly adequate computer at the time. Derman added that DEC was also adept at real-time data transmission, which is primarily what they needed.

[58] Hughes actually predicted that when a reverse salient occurs, historians will often find cases of multiple independent inventions. See Hughes, *Networks of Power*, p. 80.

[59] Tootelian interview. Other sources involved in the BASE design had never even heard of Omniswitch, or considered it inconsequential. Keep in mind that BASE was developed in a very different world: Omniswitch served members of the Interbank system, and mostly those on the East Coast.

[60] Whether this was actually technically feasible is debatable. NDC was a processor for many issuers, but not for an entire system. The computer technology of the time may not have been capable of handling all data related to the BankAmericard program, and even if it could, it would have limited the system's ability to grow, causing yet another reverse salient.

[61] 'BankAmericard Sales', *American Banker* (1 February 1973), p. 2 and 'Word BankAmericard Volume Sets Record', *American Banker* (2 May 1973), p. 3.

an automatic retry in the event of a load spike.

5.5.3 The Development Process

The process of putting together BASE requires some explanation, as it has been somewhat misrepresented in the two existing sources on Visa.[62] Here is how Hock recalled it in his autobiography:

> Swiftly, self-organization emerged.... Leaders spontaneously emerged and reemerged, none in control, but all in order.... Position became meaningless. Power over others became meaningless. Time became meaningless. Excitement about doing the impossible increased, and a community based on purpose, principle, and people arose.... A few who could not adjust to the diversity, complexity, and uncertainty wandered away. Dozens volunteered to take their place. No one articulated what was happening. No one recorded it. No one measured it. But everyone felt it, understood it, and loved it.[63]

To be fair, there are many sentiments in this quote that are supported by the testimony of others. First, those building BASE were genuinely excited by what they were doing, and many commented that they felt like they were changing the world. Second, the physical environment was also quite informal—they used hospital curtains to divide the unused space in the data center into work areas, rearranging them as needs changed.[64] Lastly, because the project had such a tight schedule and small staff, the process was a bit more chaotic than a typical TRW or IBM systems development effort.

However, Hock's account is somewhat misleading in a two important ways. First, the process was far more organized and controlled than he implies. TRW, the primary contractor, as well as the previous employer of Tootelian, made their name by applying to the civil sector the same rigorous project management techniques they developed while working on aerospace projects for the US military.[65] Those involved in the BASE project characterized it as highly-organized, though somewhat less formal due to the relatively small size of the team. Goldsmith stressed, "I've never had the privilege of being involved in such a well-organized project.... For each specialty, there was basically one cook stirring the broth. Not half a dozen. There was no committee approach to anything."

In fact, the BASE project would have never succeeded without this degree of organization, because much of the actual construction work was sub-contracted to other, off-site firms. AT&T supplied the network, DEC installed the regional concentrators, and perhaps the most critical

[62] Hock, *One From Many*; and Chutkow. Chutkow based his version largely on Hock's autobiography with some added detail from interviews.

[63] Hock, *One From Many*, pp. 173–174.

[64] Totten, Derman, Goldsmith interviews.

[65] Davis Dryer, 'The Limits of Technology Transfer: Civil Systems at TRW, 1965–1975', in: Agatha C Hughes and Thomas P Hughes, editors, *Systems, Experts, and Computers: The Systems Approach in Management and Engineering, World War II and After* (Cambridge, MA, 2000); Davis Dyer, *TRW: Pioneering Technology and Innovation Since 1900* (Boston: Harvard Business School Press, 1998).

piece, the software, was built by a vendor named Compata. The delivery schedules for all these elements had to be tightly-coordinated in order to guarantee that the system could be assembled, tested, and put into operation within the abbreviated schedule.

The primary management and coordination technique used on the BASE project was what they referred to as the "war board." A former project manager explained how it worked:

> ... in the data center..., they rented more room than they needed and they had this open area, and [George Glaser of McKinsey Consulting] put on this big flat wall all the tasks that needed to be done. It was basically a big PERT chart, but it was really more like a line of balance chart. It listed all the tasks that had to be done, and it moved left to right with time. The way he would track progress was that he had a holder for those disposable coffee cups, and he tied a string to it and used that to mark the line of balance. And he would just move the cup down the way, and as it progressed, people would report where they were in their activities. It was a very effective thing.[66]

The Program Evaluation and Review Technique (PERT), and its associated graphical representations, were originally developed for the US Navy's Polaris missile project, and it was exactly the opposite of undirected self-organization.[67] PERT was developed precisely because existing management techniques were failing to control the increasingly large and complex technical projects developed by the military after World War II. As a product of the highly-rationalized discipline of Operations Research, the PERT method promised to help beleaguered managers regain top-down control over technical projects, identify critical paths, and better predict when milestones would be completed.

The second misleading aspect of Hock's account was his use of such passive language. He seems to imply that the organizational power structure somehow dissolved and that disgruntled employees simply "wandered away" on their own accord. In reality, Hock's management style, as well as the styles of those that ran the BASE project, were much closer to a typical, authoritative director. John Totten, who compiled the requirements for BASE as a TRW contractor, and later returned to NBI as an employee, commented on Hock's account:

> ... some of the self-organization happened because people were responsible for a given task and had to fill in their progress on it. But Dee, in my judgement, never let anything just truly "float." He would say that he had all these freedoms and such, but I can remember one of his classic lines was "titles aren't important around here, as long as I'm CEO." He sometimes professed a more open management style than he really practiced.[68]

Tom Honey, who would soon join NBI and create the first debit card, echoed this:

[66] Totten interview. Note that this horizontal bar expression of a PERT model is often referred to as a Gantt chart.

[67] Willard Fazar, 'The Origin of PERT', *The Controller* (December 1962). There is some dispute as to how seriously PERT was actually used on the Polaris project, and how much it contributed to the project's ultimate success. See Harvey M Sapolsky, *The Polaris System Development: Bureaucratic and Programmatic Success in Government* (Cambridge, MA: Harvard University Press, 1972).

[68] Totten interview.

> From my experience with Dee as a new product/service developer and from observing his style with others, one was given latitude and support to do what had to be done often with unrealistic deadlines, but he made sure (and nobody questioned) who was really in charge. Rule No. 1 was that Hock was in charge and you were accountable to meet the deadline. Rule No. 2 was that if Hock was unreasonable by placing more demands on you (whether valid or invalid) that could keep you from meeting the deadline with what he wanted, you had better refer to Rule No. 1.[69]

Win Derman, who would soon be hired to manage the BASE II project, further commented that Hock had a kind of "love/hate" relationship with the technologists: "He knew he needed us, but he didn't trust us to save his life."[70] Hock had literally wagered his career by becoming the general contractor for BASE, and he certainly would not trust the outcome to an undirected, self-organized process.

The purpose of Hock's skewed portrayal becomes more clear when this account is read within the context of his entire book. Throughout the book, Hock is trying to argue that self-organizing systems (which he terms "chaordic") are actually more effective than traditional "command-and-control" organizations. A key element of his proof for this claim is the success of Visa, and therefore, it is critical that he portrays both the internal and external Visa organization as being self-organizing.[71] The development of BASE provides a tempting example, as the individual employees and contractors were given quite a large amount of autonomy in order to complete the project in such a short time. However, it is misleading to imply that the leaders were not in control of the overall project, or that the process was not highly organized.

Despite a few minor problems (such as the terminal key caps having square pegs while the Sanders keyboards had round holes), the system was completed within the $3 million budget and on time.[72] On 4 April 1973, the system was put into limited production, and by 1 May it was used twenty-four hours a day, seven days a week. After it ran successfully for a few months, NBI held a formal press conference to announce the system, which was then renamed to the more assured "BankAmericard Service Exchange."[73] However, NBI employees and member bankers would continue to refer to it by its acronym: BASE.

5.5.4 Effects of the New System

Once BASE began operation, it generated some immediate, noticeable effects. An interchange authorization that previously took four to five minutes could now be obtained in just fifty-six

[69] Personal correspondence with Tom Honey, 9 December 2006.

[70] Derman interview.

[71] By *internal*, I mean the NBI staff, and by *external*, I mean the association of members.

[72] Phillip Brooke, 'Electronic Data Transmission Net Linking All US BankAmericard Centers is Operative', *American Banker* (11 May 1973), p. 1. See also Milton O'Neil, 'Charge-Card Networks Are Working', *ABA Banking Journal* 66:3 (September 1973), pp. 116ff. Key caps story from 'BASE Is a System of People', *BankAmericard World* (June 1973), p. 2.

[73] 'BASE is Unveiled in S.F.', *BankAmericard World* (June 1973), p. 1.

seconds, and merchants could now authorize transactions twenty-four hours a day, seven days a week.[74] Although fifty-six seconds seems like an eternity from our current perspective, in 1973 this made the system *fast-enough* to be a viable competitor to cash and cheques, removing one of the critical barriers to adoption.[75]

Thresholds such as "fast-enough" or "good-enough" are an important but often overlooked dynamic in the history of technological systems. From an engineering perspective, claims are often made that one system should be chosen over another because it is faster, more reliable, etc. However, human consumers of technological systems often make their adoption choices based upon *thresholds of indifference*—once a user's experience of the system is fast-enough, it becomes an acceptable and viable choice, and it makes little difference if another system is claimed to be faster.[76] From the consumer's perspective, all systems that are faster than the threshold are equally suitable, and the customer's ultimate choice will then depend upon other factors. Of course, these thresholds will vary from group to group and will tend to "ratchet up" over time.[77] They can also be influenced by rivals who actively seek to persuade groups that their current expectations should be higher.[78]

Although there were some immediate effects, it should be noted that most of BASE's benefits were realized over a much longer period of time. After it was put into production, BASE did nothing to increase the speed of purely local authorizations—those were dependent on the acquirer/issuer's own authorization process and were never even routed to BASE. Merchants were also still operating with floor limits, so most transactions required no authorization at all.[79]

However, what BASE did provide was a *platform* upon which the system could expand and improve. Recall that a reverse salient does not necessarily prohibit the operation of the current system, nor does correcting it imply a set of immediate beneficial results. Instead, a reverse salient hampers or thwarts *further* growth, and in the case of a multi-sided platform system, an

[74] Phillip Brooke, 'Electronic Data Transmission Net Linking All US BankAmericard Centers is Operative', *American Banker* (11 May 1973), p. 1. Note that for transactions involving issuers that did not have a direct CPU interface to BASE, the average authorization time was closer to a minute and a half.

[75] Nocera makes a similar point when he says "... it was the difference between deciding to use a credit card and deciding that it was easier to use cash" (Nocera, p. 103).

[76] I first used this term in 1995 when designing a software library for accessing databases. When we needed to determine our success criteria for performance, I realized that most of our consumers would have a threshold of indifference—any library that met that threshold would be sufficient, as the library would then have little effect on the overall performance of the consumer's application. Once we reached that threshold, our development time would be better spent on other tasks.

[77] Rochlin makes a similar observation regarding consumer expectations of reliability. See Gene I Rochlin, 'Defining 'High Reliability' Organizations in Practice: A Taxonomic Prologue', in: Karlene H Roberts, editor, *New Challenges to Understanding Organizations* (New York: Macmillan, 1993), p. 17.

[78] Perception of speed is also as important as any sort of objective measurement of performance. Today, many grocery POS terminals will encourage the cardholder to swipe their card while the checker is still adding items to the total. By doing so, the POS terminal can capture the card information, prepare an authorization message in memory, and open a connection to the acquirer, all while the clerk finishes dragging your items over the scanner. By overlapping these tasks, the consumer perceives the overall authorization time to be shorter.

[79] The merchant was required to check the number against the warning postcards, but as noted in earlier chapters, this was commonly ignored.

inability to grow and adapt could actually cause a system collapse.[80]

Specifically, BASE enabled NBI to accomplish four things over the next few years. First, NBI could now expand the system by recruiting additional member banks without compounding the problems surrounding manual interchange authorization. Although only a small percentage of transactions required interchange authorization in 1973, that small percentage would continue to grow in absolute terms as more members joined the association. Second, NBI could also now encourage consumers to travel with their card, using it in non-local settings, which would naturally increase the percentage of those transactions requiring interchange authorization. This was attractive to NBI not only because it made the card more useful, but also because an increase in interchange authorizations would result in an increase in fees paid by the members.[81] Third, as more member banks automated their local authorizations, NBI could begin to reduce floor limits, which would help reduce fraud and bad credit charge-offs.[82] Lastly, and most importantly, BASE created an online computer network connecting all the member processing centers, and this network made it easier to automate other interchange-related functions.

5.6 Conclusion

In this chapter, we saw how the authorization process had become a reverse salient for all the bankcard systems, and NBI in particular. To correct this salient, the industry leaders formulated two critical problems: the local authorization decision was too slow because it required human intervention; and interchange authorizations were too slow because they required manual "two-legged" calls or telexes. The solution implied by the first problem was the replacement of human authorizers with automated, computerized logic, and we reviewed the various ways in which organizations accomplished this in the early 1970s. The solution implied by the second problem was the development of high-speed electronic communication between the card centers, and we discussed how both Omniswitch and NBI developed similar but independent switching networks.

In the next chapter we will turn our focus to NBI's solution to the second operational problem discussed in the previous chapter: clearing and settlement of interchange transactions.

[80] Evans and Schmalensee, pp. 131–158.

[81] This meant more operating revenue for NBI, which would enable them to automate more functions and expand their processing capability.

[82] Floor limits were not entirely removed until after the development and mass-adoption of inexpensive dial-up POS terminals, which I will discuss in Chapter 8.

Chapter 6

Automating Clearing and Settlement: BASE II and III

In the previous chapter, we explained how NBI and others automated both local and interchange authorizations. The development of the BankAmericard Service Exchange (BASE) provided NBI with a powerful platform that enabled not only the switching of interchange authorizations, but also any kind of online message exchange between BankAmericard processing centers.

In this chapter, we will discuss NBI's next major computer system development projects, known as BASE II and III. These systems were designed to automate the clearing and settlement of interchange transactions, which was the second major operational problem area discussed in Chapter 4. BASE II replaced the cumbersome mailing of paper drafts between members with a centralized, batch-oriented, electronic clearinghouse. BASE III was intended to replace the old BofA punched-card accounting system given to new members with a modern, sophisticated mainframe computer program capable of seamless integration with BASE I and II. However, building generic software for the member banks' processing centers was not only a significant departure from NBI's core purpose, but also an activity they did not entirely understand. As a result, the BASE III system failed to meet its objectives, and was eventually cancelled. However, as we shall see, Hock ultimately managed to turn this first serious defeat into a personal victory that helped to ensure the financial success of the organization.

6.1 Truncating the Paper: BASE II

In Chapter 4, I described the manual, tedious, and error-prone process for clearing and settling interchange transactions in 1968. Although the BankAmericard organization had been restructured in 1970, and the switching of interchange authorizations had been automated in 1973, there was still no centralized clearinghouse. Acquiring banks continued to sort and mail the physical sales drafts to each issuing bank, where they were reconciled against clearing drafts that had already been received and paid separately.

The member banks were able to cope with this process while the volumes remained low, but as more banks joined NBI, and the sales volume increased by thirty to forty percent each year, so did the number of interchange transactions. During 1972, the NBI member banks exchanged 95 million drafts, and they projected that this would rise to 225 million by the end of 1975.[1] It was clear that without an automated, centralized clearinghouse, the BankAmericard system would grind to a halt.

As noted in the previous chapter, automating interchange authorizations was just the first phase of Hock's overall plan to build an electronic value exchange system. Once BASE was put into operation in 1973, Hock quickly began the second phase, which was aptly named BASE II.[2] In this phase, he intended to automate the clearing and settlement of interchange transactions, but instead of using high-speed MICR readers and sorters as the Fed had done with cheques, he wanted to *truncate* the paper, transforming the sales drafts into electronic records, and clearing them through a centralized computer system.[3] Using this approach, interchange transactions could be cleared and settled as early as the night after they were deposited.[4]

However, building an electronic clearing system would require engineering more than just computers and telecommunications. It also required the "engineering" of a more cultural dynamic: the consumer preference for country-club billing.[5]

6.1.1 Country-Club, Descriptive, and Facsimile Billing

In 1973, most NBI member banks performed *country-club billing*, which is the practice of returning the punch-card layer of each sales draft to the cardholder, along with a summarized bill. The term originates from the traditional practice of country clubs in the US, nearly all of which would bill their members by returning the "chits" signed when charging purchases at the club.[6] However, for most cardholders, this billing method was not as reminiscent of the genteel country club as much as it was of their monthly current account statements. At this time, most banks returned cancelled cheques to account holders, providing them with not only a visual memory of the transaction, but also a legal proof of payment. Thus, it seemed "natural" to return bankcard sales drafts as well.

[1] 'NBI Planning Paperless Card Drafts', *American Banker* (18 December 1973), p. 1.

[2] The original working name for BASE II was actually "Shared Paperless Activity Network" or the "SPAN between the banks" (Derman interview). "Paperless" commonly appeared in the names of electronic clearing systems at this time, highlighting that the primary goal was to eliminate the flow of paper.

[3] Note that with the passage of the "Check 21" Act in October 2003, cheques in the US may now be truncated and cleared electronically. See Will Wade, 'President to Sign Check 21 Today', *American Banker* (28 October 2003), p. 23. In fact, some merchants now simply pass cheques through a scanner and hand the physical paper back to the customer.

[4] BASE II could clear transactions overnight, but at this time, the acquirer often needed a day or more to convert the draft into electronic form. See section 6.1.2 on page 106.

[5] My language here draws upon John Law's idea of "heterogenous engineering." Law used this term to describe the process by which system builders attempt to enroll both human and non-human actors into their networks. He notes that system builders often need to engineer social and political conditions in addition to physical or technical ones. See Law.

[6] Jutilla, p. 181.

Interview sources noted several reasons why many cardholders preferred country-club billing. First, the drafts provided them with instant visual memories of each transaction, as merchants occasionally wrote a short description of the purchased merchandise in the detail area. Second, cardholders could also quickly verify their signatures on each draft to ensure that each was a legitimate charge. Third, those who kept the customer receipt layers from the time of sale could easily compare those against the punch-card layers returned by the issuer, ensuring that amounts had not been altered, and noting which transactions were still pending. Fourth, for those who did not wish to keep the customer receipts, the punch-cards returned by the issuer could be used as an evidence of payment for expense reports and income tax returns. Technically, bankcard sales drafts were not legal proofs of payment like cancelled cheques, but they were effectively treated as such by cardholders, their employers, and most importantly, the US Internal Revenue Service.[7]

Country-club billing was also advantageous for smaller issuers. Because the drafts were punch-cards, they could be sorted, tabulated and collated by rather simple mechanical devices, instead of expensive mainframe computers.[8] Summary bills could be generated using the old BofA software that was still given to new members when they joined NBI. Additionally, because the drafts were returned to the cardholder, the issuer did not need to capture the merchant names and locations in electronic form, as they were printed directly onto the drafts.[9]

An alternative to country-club billing, called *descriptive billing*, began to appear in the late 1960s, but was not widely used in the BankAmericard system until after BASE II was completed.[10] In this method, the issuer captured the information from the sales drafts in electronic form, stored it in a computer file, and sent cardholders only a simple list of their transactions.[11] For each transaction, the statement listed the purchase or posting date, the merchant name and location, and the amount. Because the transaction information was reproduced in descriptive form, the physical drafts were not returned to the cardholder.[12] The issuer stored the drafts for a short period of time, and then maintained a microfilm copy in case the cardholder disputed the charge.

The descriptive billing method was preferred by issuers with larger volumes because it

[7] See the "Supporting Documents" section of IRS publication number 583, `http://www.irs.gov/publications/p583/ar02.html` (accessed on 18 December 2006). The difference between "proof" and "evidence" of payment is subtle—the former is a legal proof recognized by the courts, and supported by the Universal Commercial Code, while the latter is something that a particular party (e.g., an employer or the IRS) would accept as sufficient evidence.

[8] Dumler interview.

[9] The details of how bankcard centers accomplished data capture will be discussed in section 6.1.2 on page 106.

[10] For an explanation of descriptive billing and its benefits for the banks, see Nicholas E Magnis, 'Descriptive Billing—A Solution to the Paperwork Problem', *ABA Banking Journal* 62:10 (April 1970).

[11] Before BASE II was operational, the acquirers still mailed the physical interchange drafts to the issuer, so it was the issuer who performed the electronic data capture for descriptive billing. As we shall see, BASE II shifted this responsibility to the acquirer. However, keep in mind that the majority of transactions at this time were still local or "on-us," meaning that the acquirer and issuer were the same bank.

[12] Some banks used a hybrid method, returning the physical drafts for on-us transactions, but providing a descriptive bill for interchange transactions (Dumler interview).

required handling the paper only once, which greatly reduced their labor and postage costs. It was also preferred by NBI, because it did not require the movement of paper from acquirers to issuers. Ultimately, NBI wanted to truncate the paper at the acquirer and clear the transactions in electronic form, but this would be impossible if cardholders continued to demand the original drafts.

However, neither Hock nor his staff at NBI believed that all cardholders were actually "demanding" the original drafts; they were merely *accustomed* to receiving them. Although some cardholders were no doubt reassured by the original, most would probably not notice if they received something that merely *looked like* the original, as long as it contained the most important information: the merchant name and location; the date of the transaction; and the amount. In fact, cardholders already received on occasion a clone of the original draft if it was badly mangled by the merchant or the punch-card readers.[13] If the issuer could print what NBI called a *facsimile draft*, based upon transaction information electronically transmitted from the acquirer, most customers might not even notice the difference, and even if they did, the facsimile would still be an adequate evidence of payment. If the cardholders accepted the facsimile drafts without serious complaint, the bank was then one step closer to implementing descriptive billing.

To test this hypothesis, NBI conducted an experiment with six of its more technically-advanced banks starting in early 1973.[14] Acquirers of interchange drafts captured the descriptive billing information in electronic form, and then transmitted it to the issuer. The issuer then computer-printed the transaction information onto the punch card layer of a new, blank sales draft and sent that to the cardholder with their summarized bill.[15] The facsimile looked nearly the same as the original, but did not contain any details of the purchased merchandise, nor the cardholder's signature. The lack of item detail actually posed little problem as most merchants neglected to write anything meaningful, and as long as the customer recognized the charge, their signature was superfluous. If the customer could not recognize the charge, or needed the original for some other purpose, the customer could still request it from their issuer. The issuer would then request it from the acquirer, who would send it through the mail. Effectively, the movement and return of physical sales drafts became the exception instead of the rule.

In October 1973, NBI announced that "neither its member banks nor cardholders have related any major problems or objections to the facsimile drafts. . . ."[16] Again, it was not that cardholders were *delighted by* the facsimile drafts; they simply did not care enough to complain about them to any significant extent. After a few months, they also became accustomed to them, and requests for originals were infrequent. During the test, cardholders requested only

[13] Dumler interview.

[14] 'NBI Finds No Major Problems With Facsimile Drafts', *American Banker* (10 October 1973), p. 81. Confirmed in Dumler, Derman, and Russell interviews.

[15] Note that this was not an exact, pixel-for-pixel copy of the original draft. The new draft was printed from discrete, alpha-numeric data.

[16] 'NBI Finds No Major Problems'.

1 original for every 500 facsimiles. Forty-five percent of those were due to the cardholder not recognizing the acquirer's processing name for the merchant, which often differed from the merchant's trade name, but this could easily be adjusted. Nineteen percent were requested for business purposes, typically for employers who were wary of the facsimiles, and those cardholders quickly learned to keep the customer receipt layer of the original draft. The remaining seventeen percent were requested by the issuer for fraud analysis and prosecution.[17]

In essence, this test was the key to determining if BASE II was *culturally* possible. NBI could engineer the computer systems and telecommunication networks, but if they could not also "engineer" the cardholder acceptance of facsimile drafts (which were just descriptive billing in a more recognizable form), cardholders would have demanded so many of their originals as to negate the benefits of an electronic clearing and settlement system. As consumers used the card for more purchases, and descriptive billing became the norm, some cardholders did begin to complain, and in 1977, Rep. Frank Annunzio (D-Ill) proposed a bill that would have effectively outlawed descriptive billing.[18] However, this section was removed from the final bill, as it was then too late to mandate a return to country-club billing, and the committee eventually realized that ensuring cardholder rights during charge disputes was the real issue. However, the 1973 test showed that facsimile drafts would be acceptable to the current cardholder base, allowing NBI to continue with the design and implementation of BASE II.

6.1.2 Design of BASE II

The design of BASE II fell to a newly-hired employee named Win Derman.[19] A graduate of MIT and Stanford, Derman became NBI's twenty-fifth employee after a stint with the Stanford Research Institute, where he had specialized in management engineering and systems design. B Ray Traweek was also hired to manage the overall project. Traweek's extensive experience with computer systems began while earning a Masters in Mathematics from the University of Texas, and was honed while working for Convair (now part of General Dynamics) and TRW. While at TRW, he had worked on a banking system with Aram Tootelian, and when the BASE II project began, Tootelian convinced Traweek to come to NBI to lead it.

NBI had originally planned to use the same network of Sanders terminals developed for BASE (now renamed BASE I to avoid confusion) for the electronic clearing and settlement of interchange transactions. They assumed that acquirers could manually key the transactions into the terminal, sending each as a new type of message to the central switching computer.

[17] A later story from December quotes Hock as saying that 1 in 300 were requested, so either the ratio increased somewhat over the two-month period, or it was reported incorrectly. See 'NBI Planning Paperless Card Drafts', *American Banker* (18 December 1973), p. 1. Either way, the ratio was still small enough to justify building BASE II.

[18] The text of the original bill, HR 8753, and a transcript of the hearings is available in 'Consumer Credit Protection Act Ammendments'. See also, 'Annunzio Accuses Credit Card Companies of "Consumer Entrapment"', *The Nilson Report* 169, p. 1.

[19] Details of the BASE II project primarily come from interviews with Derman, Traweek, Russell, Goldsmith, and Peirce.

However, Derman quickly realized that the growing number of transactions would ultimately make this approach impractical, so he convinced the NBI management that BASE II should use a batch transmission approach instead.

In essence, the new BASE II design was a computerized version of the clearinghouse concept discussed in Chapter 3. Although BASE II was not the first implementation of this idea, commonly called an *automated clearinghouse* (ACH), it was by far the largest and most ambitious, the first with a national scope, and the first in the domain of bankcards.[20] With the BASE II system, NBI would act as the centralized clearinghouse of all BankAmericard interchange transactions, but instead of exchanging paper, the acquirers and issuers would exchange electronic records of those transactions. Instead of maintaining high-speed MICR readers and sorters, NBI would maintain one central mainframe computer to collect, sort, total, and distribute the transaction data. Furthermore, because the electronic transactions could be transmitted over telecommunication lines, all the BankAmericard processing centers in the entire nation could clear and settle through the clearinghouse every night.

Additionally, members would now also settle *with the clearinghouse* instead of each other. Prior to BASE II, the BankAmericard system performed *bilateral gross settlement*, meaning that each acquiring bank B1 collected from each issuing bank B2 the total amount (less interchange fees) of all transactions involving the issuer's cards. However, since nearly every acquirer was also an issuer at this time, B2 might also collect roughly the same amount from B1 on the same day. In theory, this could result in $n(n-1)$ transfers of similar amounts between every pair of members every day. With BASE II, each member would settle only with the clearinghouse, resulting in just one value transfer per member per day. Furthermore, BASE II would perform *net settlement*, meaning that the amount each member owed the clearinghouse would be subtracted from the amount the clearinghouse owed the member, and the member would pay or receive only the difference.[21]

Data Capture

However, in order for all of this to work, the acquirers needed to encode the paper sales drafts into an electronic form that could be transmitted to NBI's data center. This process, known as *data capture*, could be accomplished either by manual entry, or by scanning the drafts using an Optical Character Recognition (OCR) device. The latter promised to be faster and more accurate than the former, but even in the ideal case, the OCR scanners of the early 1970s could

[20] In the late 1960s, the California commercial banks organized the Special Committee on Paperless Entries (SCOPE), which resulted in the creation of the California Automated Clearinghouse Association (CACHA) (Yeatrakas interview). This organization, with the help of the Federal Reserve Bank of San Francisco, began operating what seems to be the first ACH in the US on 13 October 1972. See Phillip Brooke, 'In Past Year, EFTS Concepts Shift from Theory to Action', *American Banker* (4 June 1973), p. 13. However, this was only a regional system for direct deposit and pre-authorized debit transactions, whereas BASE II was a national system for bankcard transactions.

[21] Although this was new for the BankAmericard system, most other clearinghouses already used the net settlement technique to reduce the amount of funds transfered.

capture only half of the information.[22] The card and merchant numbers were printed in a standard OCR font, but the transaction date and amount were typically hand-written on the draft by the merchant. Eventually imprinters with adjustable embossed wheels for the date and amount became available, allowing the OCR scanners to read the entire draft automatically. However, these imprinters were not widely adopted, as they were more expensive and consumed more counter space.[23] Additionally, the hype surrounding Electronic Funds Transfer Systems (EFTS) promised that all paper would soon be eliminated from all aspects of banking, so acquirers were not concentrating on how to make the processing of paper drafts more efficient.[24] In fact, many of the largest acquirers continued to use manual data capture until the widespread adoption of merchant point of sale terminals in the mid 1980s.

Although moving to a system of electronic data capture required an enormous amount of work and expense on the part of acquirers, they were, in many ways, eager to change. This is how Derman explained it:

> The reason that the member banks turned heaven and earth to do BASE II can be exemplified by one comment: a banker in New Jersey took me out in the hall and said, "look, you're going to make me do a tremendous amount of work to change... but I'm going to do it. And you know why? Look up there on the wall; see that line? That's how high we stack the sales drafts that come in from our merchant banks when they process their Christmas volume, which is when we do half our business. And it takes us *four months* to work through that! And we're at risk for all those transactions and all that fraud during that four month period."[25]

Recall that acquirers credited a merchant's account upon deposit of the drafts, but the acquirer could not recover those funds until the drafts were processed, either by billing the cardholder for local transactions, or clearing and settling the non-local drafts through interchange. Any delays in processing resulted in increased float, and corresponding costs for the acquirer. Additionally, issuers were liable for any fraudulent transactions, and had no way to stop further purchases until they processed the drafts and detected that the card had been compromised. Thus, by shifting to semi- or fully-automated OCR data capture, acquirers could not only process their local transactions more quickly, they could also clear and settle their interchange transactions the same night through BASE II.

[22] 'What Credit Card Executives Should Know About OCR Readers', *The Nilson Report* 174, p. 1.

[23] Counter space is actually an often overlooked, yet critically-important dynamic in the history of payment systems. Many merchants have very limited counter space, and any system that requires a separate, large countertop device faces a significant barrier to adoption.

[24] Nilson rebuked this approach as myopic: "At this point in time, handling credit/debit transactions via electronic terminals in any volume is only a mirage. Except for automated teller machines and cash dispensers, EFTS is a bust!... Banks, which until recent months had hoped EFTS would reduce the paper flow, must now find other ways to deal with the volume of sales slips which will increase at least 100% in the next five years as predicted in my last issue" (*The Nilson Report* 174, p. 1).

[25] Derman interview.

Edit Package and TTUs

After acquirers captured the draft information in electronic form, they could then easily separate interchange transactions from local ones. However, they were no longer required to sort and total the interchange transactions by issuer. Instead they simply submitted them in batch to the central clearing computer, which would sort and total all transactions submitted from all acquirers.

To get the transactions to the BASE II central computer, acquirers first needed to transform their data records into the BASE II format and validate them. The acquirers wrote their own software to convert formats, but validation was accomplished by running a program supplied by NBI, called the *edit package*.[26] This program ensured that all data were present, dates were in the correct format, card numbers were valid, and anything else that was necessary to ensure that all the transactions in the batch could be properly cleared and settled. The program also produced a number of reports that the member banks could use as a printed record of their outgoing transactions.

Although it was costly for NBI to maintain a version of the edit package for each kind of computer used by the processing centers, it was ultimately advantageous because they could clear and settle batches of transactions without having to handle individual exceptions. If the batch passed the edit package, every transaction could be cleared; if it did not, the acquirer had to correct the problematic transactions, or remove them from the batch. By making the batch the atomic unit, NBI greatly simplified the task of interchange accounting, as the amount submitted to interchange would always equal the amount credited, minus fees.[27]

After validating the interchange transactions using the edit package, acquirers copied them onto a magnetic tape, which was then mounted onto their *tape transmission unit* (TTU). The TTUs were custom-engineered DEC PDP-11/10 minicomputers equipped with a tape drive, modem, and a bell (explained later).[28] NBI contracted with DEC to install and maintain one of these devices in each of the eighty-eight BankAmericard processing centers in the US.

Central Clearing Computer

At the core of the BASE II system was a large mainframe computer. Each night, starting at 5:00 PM Pacific Time, the central computer would begin calling each of the TTUs according to a schedule established with the processing centers. Because the TTU was a minicomputer, it could answer the call and establish communications automatically. For five hours, known as the *input phase*, the central computer initiated connections to each TTU and read all the transactions contained on each tape. During the validation process, the edit package inserted

[26] Derman and Traweek interviews.

[27] Reportedly, Interbank's automated clearing and settlement system initially allowed individual exceptions, but they quickly moved to NBI's model as the exceptions were too difficult to manage (Derman and Powar interviews).

[28] Derman interview.

verification amounts onto the tape so that the mainframe program could ensure that the information was transmitted accurately; if it had not, the TTU could automatically backup to the last checkpoint and resume transmission.[29] At the end of a successful transmission, the central mainframe then sent an instruction to the TTU that triggered it to ring the aforementioned bell. This bell told the operator in the processing center to unmount the tape containing the outgoing transactions, and mount a new blank tape for the incoming transactions and summary reports.[30]

The central computer then moved into a two hour *sort and calculation phase*, during which it performed a number of tasks necessary for clearing and settlement. It first calculated the amount owed to each acquirer, which was the total of their submitted transactions, minus the interchange reimbursement and NBI's processing fees (1.95 percent and 2.5 cents per item respectively). It then sorted the transactions by issuer and calculated how much each issuer owed. Finally, it computed the net amount each member owed or was due from the clearinghouse. This information was then printed by NBI for use in the actual settlement, as well as auditing and accounting.

The final phase of the BASE II cycle was called the *output phase*. For five more hours, the mainframe again established connections with each of the TTUs, and streamed back all the interchange transactions for which that bank was the issuer. The member banks could then extract these transactions using the edit package, and incorporate them into their own billing systems. In addition to the incoming transactions, the central system also transmitted a full clearing report (which the member could use for bookkeeping and reconciliation), as well as the net settlement amount for that bank. By 5:00 AM Pacific Time the next morning, each bank had all the information they needed to bill their cardholders and settle with the clearinghouse.

The actual movement of "good and final funds" was still accomplished with clearing drafts, but now NBI completed these drafts on behalf of the members. After the clearing drafts were prepared, an NBI employee literally got in her car and drove them to a BofA branch located just down the hill from NBI's data center.[31] By 1980, the settlement process was also automated by transmitting the net settlement amounts electronically to a clearing bank.

Although 5:00 PM seems like an appropriate "close of business" time to begin the BASE II process, Derman explained that this start time was actually chosen for a more significant reason:

> We, even in those days, thought this was a worldwide system, so we ran everything worldwide. We didn't arbitrarily pick 5:00 PM. We said 5:00 PM is GMT 0:00.[32]

Although the original BASE II provided nightly clearing and settlement for the US member

[29] Traweek interview. Confirmed in Phillip Brooke, 'BankAmericard Begins Electronic Interchange of Drafts', *American Banker* (6 November 1974), p. 1.

[30] Derman interview. Unfortunately, the tapes at this time did not contain enough capacity for both the outgoing and incoming transactions.

[31] Kollmann, Nordemann, and Harrison interviews.

[32] Derman interview. Note that GMT does not adjust for Daylight Savings Time, so the local start time for BASE II would be different during the summer.

banks only, as we shall see, it was eventually expanded to be the clearinghouse for all members worldwide.

Making Up With IBM

The original capacity target for BASE II was to clear one million transactions within their allotted twelve-hour processing window.[33] Considering the rate at which the system volume was increasing, it was also likely that they would need to expand their capacity within just a few years. Unfortunately for Hock, the only mainframe computer capable of processing that kind of load and enabling seamless capacity upgrades was made by IBM. Recall that Hock had previously sworn never to do business with IBM again, but the capacity requirements for BASE II made that pledge difficult to maintain.

Fortunately, IBM had recently transferred the NBI account from the Banking to the Manufacturing and Distribution Office, and it was now in the capable hands of an IBM sales representative named Roger Peirce. Peirce had been with IBM since 1963, first as a systems engineer and later as a direct commissioned salesperson. He had all the qualities that Hock respected: a sharp intellect, technical competence, business savvy, and a direct, no-nonsense communication style. Additionally, while developing numerous information systems, Peirce had learned not only what was required for a successful project, but also how to handle difficult customers.

Traweek and Derman knew that BASE II required an IBM mainframe, but convincing Hock to buy one would be a challenge considering that Hock typically refused to talk with IBM sales representatives. They approached Peirce and asked him to do something to soften Hock's temper. Peirce explained what they did:

> We arranged what we called "the pillow call." ... We got the highest ranking guy we could find in IBM... and we convinced him to come out. We said "look, this is going to be an unpleasant call—strap a pillow on your ass because you're going to get beaten!" So we went into this meeting and Dee basically railed and ranted at the guy for an hour and then went out, but after it was all done, it worked, and they decided they would do business with IBM.[34]

NBI ordered a System/370 model 145, and Peirce called in a number of favors to advance its delivery so that it could be installed in time.[35]

Peirce personally oversaw the installation of the mainframe in April 1974, which no doubt endeared him to Hock. IBM, the most important computer company in the world, was now treating Hock as the important client he saw himself to be. As a result, Hock became more amenable to IBM, and Peirce in particular, which would eventually enable IBM to sell Hock

[33] Derman interview.

[34] Peirce interview.

[35] Peirce interview. This was the first IBM mainframe to use silicon memory chips instead of a magnetic core. See `http://www-03.ibm.com/ibm/history/exhibits/mainframe/mainframe_PP3145.html` (accessed on 15 December 2006).

on moving BASE I to their platform as well.[36] Peirce would also later join NBI and become Vice President of Systems Development, Operations, and Member Relations.

6.1.3 Final Development and Rollout

In addition to IBM, NBI contracted with a number of other vendors to construct the various pieces of BASE II. DEC built, installed, tested, and maintained the TTUs at the eighty-eight BankAmericard processing centers. AT&T again supplied the telecommunications. Arthur Andersen designed the audit control system, and the Stanford Research Institute (SRI) performed the technical acceptance tests.[37] Compata, the organization that had written the BASE software, also returned to write the software for BASE II.[38] Using the same project management techniques they used for BASE, NBI completed the new system by November 1974, within its allocated budget of $7 million, and timeframe of 18 months.[39]

However, some of the processing centers still had not yet implemented their data capture systems by November. Thus all could *receive* interchange transactions from BASE II, but not all could *send* them. Therefore, NBI mandated that on 1 November 1974, every issuer had to receive electronic interchange transactions via BASE II, but the acquirers who had not yet automated could continue to mail paper drafts through the Christmas season; by 1 March 1975, all interchange transactions had to be cleared electronically. Both deadlines were met by the members, and the mailing of original drafts became the exception rather than the rule for the entire system.[40]

6.1.4 Effects of the System

BASE II altered the BankAmericard system in five important ways. First, BASE II dramatically reduced the time it took to clear and settle interchange transactions. Under the manual system, it took an average of six to eight days for sales drafts to reach the issuer, where they often failed to reconcile with the already paid clearing draft; with BASE II, all sales drafts were now cleared and settled, in batch, overnight.[41]

Second, the reduction in clearing time resulted in a corresponding reduction in float, mostly for the acquirers, but also for the issuers. BASE II essentially forced the acquirers to implement an automated data capture system, enabling them to submit transactions to interchange much

[36] Details of the BASE I port to IBM and the TPF operating system will be discussed in the next chapter.

[37] Although SRI had already separated from Stanford University in 1970, they did not formally change their name to SRI International until 1977. See `http://www.sri.com/about/facts.html`.

[38] Brooke, 'BankAmericard Begins Electronic Interchange'.

[39] Hock, 'Electronic Funds Transfer or Electronic Value Exchange?', p. 13.

[40] It is important to note that on this latter date, the paper did not entirely disappear from the system. The paper was eliminated for the most part from interchange, but most merchants still completed and deposited paper drafts, and acquirers still handled that paper during data capture. The paper was not removed from this segment of the clearing path until Visa coordinated the development of small, inexpensive point of sale dial terminals. That story will be discussed in Chapter 8.

[41] Brooke, 'BankAmericard Begins Electronic Interchange'.

more quickly, and thus recover the funds they had already credited to their merchants. Issuers received the transaction details electronically at the same time they paid the acquirers, which not only eliminated the painful reconciliation process, but also provided an easy way to import the transactions into their billing systems in order to recover funds from the cardholder.

Third, automating the interchange process dramatically reduced the labor and postage costs associated with the clearing and settlement of interchange transactions. NBI estimated at the time that BASE II saved the members between $14 and $17 million in gross clearing costs during its first year of operation alone.[42] Although the exact amount would have been difficult to substantiate, as so many aspects of the business were in constant flux, there is no question that the labor costs alone would have made the existing manual process uneconomical as the number of interchange transactions increased.

Fourth, faster clearing also meant that issuers now received fraudulent interchange transactions in a more timely manner. Although BASE I stopped many fraudulent transactions, it only saw those that were over the merchant's floor limit. A compromised card number could still be used for numerous charges under the floor limit. The issuer would have no way of knowing that fraud was occurring until those transactions were cleared through interchange. The faster transactions were cleared, the faster the issuer could detect fraud, and take steps to cancel and recover the card.

Lastly, BASE II established a platform upon which NBI could offer other *batch-oriented* data transfer services between the members, just as BASE I provided a platform for *online* message exchange.[43] From its inception, BASE II allowed members to transmit other administrative transactions such as chargebacks, reversals, and requests for originals, in addition to interchange drafts.[44] BASE II was later extended to include other kinds of clearing messages as well, such as rewards for recovering stolen cards, or reimbursements for telex costs incurred by foreign acquirers.[45] With the combination of BASE I and II, NBI could now facilitate any type of data exchange between its members. These two systems became the information processing backbone upon which Hock could eventually provide his "premier system for the exchange of value."

6.2 Losing Focus: BASE III

Shortly after NBI began the BASE II project, they also embarked on another system that has, so far, received almost no attention, perhaps because it was the organization's first significant

[42] Brooke, 'BankAmericard Begins Electronic Interchange'.

[43] The term "online" is often used in many different ways. Here I mean a system that maintains constant communication links between nodes, passing individual messages in a near-real-time manner. Some news accounts referred to BASE II as being "online" because it transmitted data over communication lines (as opposed to mailing magnetic tapes), but it is more appropriately described as a batch-oriented data exchange system.

[44] Brooke, 'BankAmericard Begins Electronic Interchange'.

[45] Kollmann interview.

failure.[46] While BASE I and II enabled data exchange *between* the member processing centers, this system, appropriately named BASE III, was designed to provide most of the necessary card-related data processing *within* a given center. It would replace the simple punched-card system originally developed by the BofA with a modern, on-line mainframe computer program that would provide all the functionality necessary for maintaining cardholder accounts. In essence, it was to provide "everything you needed to plug into Visa."[47]

NBI had two primary motivations for building such a system. First, in keeping with the general philosophy of networked payment card systems, NBI thought that if all members co-operated, they could build a better system than any one of the members could have built alone. At the time, Hock assumed that all the processing centers had more or less similar needs, so it made little sense to build eighty-eight versions of the same software. NBI felt that they could manage the construction of a much more sophisticated, shared computer program for less money by combining all the member's resources. Once it was finished, those members that contributed the initial funding would receive the software, and as more members chose to adopt it, the original funders would receive a corresponding portion of the subsequent licensing fees.[48]

Second, if NBI developed a shared computer program, they could also enable a consistent and seamless integration with the BASE I and II systems. Recall that NBI wanted to enable fully automated authorizations by connecting BASE I directly to each issuer's computer, and BASE III would provide them a perfect opportunity to deploy this functionality to a large number of processing centers. Additionally, once BASE II began sending incoming interchange transactions, BASE III could include the software necessary to import those easily into the cardholder files.[49]

NBI began this effort in mid 1973 with the ambitious intention of releasing it shortly before BASE II was put into production (November 1974). As opposed to their experience with the original BASE project, this time they found some existing software that they assumed could be expanded to meet the needs of all processing centers. This had been developed by the Centurex corporation of Los Angeles for the American Bank and Trust of Reading, Pennsylvania, which also happened to be an NBI member.[50] NBI hired Centurex in July of 1973 to help gather requirements from eighty other processing centers and extend their existing system accordingly.

[46] Hock's autobiography is the only published source that contains any mention of BASE III. See Hock, *One From Many*, pp. 241–243.

[47] Totten interview. Information on BASE III comes primarily from interviews with Totten, who managed the project in its later phases, as well as Fojtik, Peirce, and Russell.

[48] Hock, *One From Many*, p. 242.

[49] Totten interview.

[50] Phillip Brooke, 'NBI Hires Centurex to Design Software System for Member Banks', *American Banker* (1 August 1973), p. 1. See also 'NBI Buys Centurex On-Line Card System', *American Banker* (2 January 1974), p. 5.

6.2.1 Problems Begin to Emerge

It was during the requirements gathering phase of the project that the first problems began to emerge. Contrary to assumptions, the processing centers did have varying needs, and the larger centers required specific features that necessitated significant changes to the underlying data structures and logic. The most serious was what became known as the "multi-bank" feature—centers that processed for multiple member banks (e.g., the BofA) required the system to segregate cardholders and transactions according to issuing bank, something that the original single-bank system had never anticipated.

However, instead of partitioning these more advanced requirements into processor-specific extensions, those gathering the requirements promised to include all these features in the main software. John Totten, who joined NBI in February of 1974 to lead the BASE III project, explained:

> Part of the problem with that was, many of the people that were on the project had never been in a credit card [processing] environment before, and so they would have these road shows and somebody would say "will it do X, will it do Y, will it jump over this and crawl under that?" And the typical answer was "yeah, if you sign up, we'll put that in."[51]

Because BASE III was entirely funded by license fees paid by advance subscribers, there was strong incentive to agree to any feature necessary to secure the sale. Members also had an incentive to subscribe early as the license fee was less for advance subscribers, and according to Hock, they would also receive a portion of the fees paid by future subscribers. As a result, the list of requirements grew far beyond what could reasonably be accomplished by a single, shared program to be delivered in November 1974.

In addition to a growing requirements list, the system was also being designed to the lowest common denominator of computer hardware. For example, many of the smaller centers were still using the simpler IBM 2260 video display terminals to interact with their mainframes, and were unwilling to upgrade to the more powerful, but expensive 3270. Thus BASE III was constrained to use the 2260 even though the larger and more innovative centers had already upgraded to the 3270. Roger Peirce noted that because NBI had little experience with the IBM environment, they did not yet understand how much of a constraint this would be. He described the BASE III design approach as "skating to where the puck is instead of where the puck is going to be."[52]

By the time Totten took control of the project, NBI had already spent all the money raised by selling advance licenses, and were still ten months away from their planned first release.[53] Furthermore, the first release was simply a more generic form of the original Centurex software. Most of the new features would not be available until the second release, which was planned

[51] Totten interview.
[52] Peirce interview.
[53] Totten interview.

for ninety days after the first, and the multi-bank capability would require yet another release, which was promised for ninety days after that. The project was clearly in trouble, and Hock was forced to ask the Board for additional funds. Hock later admitted that asking for more money was in actuality an unwillingness to recognize and admit failure, and only served to further compound the problems.[54]

BASE III encountered yet another problem during the development phase that is common in rapid parallel development. BASE III was supposed to import incoming interchange transactions from BASE II into the cardholder files, but BASE II itself was being developed at the same time and thus became a moving target. Totten explained:

> I can remember, my development team was in Los Angeles, and we'd be working on some sort of details about the transaction format, and I would walk down the hall and Win [Derman] or B Ray [Traweek] would say, "oh, we just changed the transaction format." I asked them, "How in the world am I going to have a system that's compatible with you if you guys are continually making these changes, which I have to accommodate, and I learn about it over the water fountain?"[55]

6.2.2 Problems Come to a Head

Centurex did eventually deliver the first version of BASE III in late 1974, but few centers chose to install it, as the new features were still to come ninety days later in the second version. When that second version became available, those that installed it found the performance to be somewhat less than satisfactory. Russell explained:

> ...we purchased a rotten piece of software, and proceeded to "improve" it. As you well know, improving rotten software only exacerbates the rotten aspects of it, and when we were finished we had a beautiful piece of garbage which had but one flaw—it required 25 hours to run! Now, even the genius of Dee couldn't figure out how to add another hour to every day, and so it failed, miserably. Unfortunately, 4 or 5 banks installed this software, and at least 3 of those plan managers were fired for their decisions.[56]

Unfortunately, NBI had promised to deliver a third version of BASE III in another ninety days, and this version would require substantial changes and additional code to accommodate the multi-bank feature, which would likely make the system even slower and less reliable. A number of key players including Traweek, Totten, and Peirce convinced Hock that BASE III was doomed, and that NBI should admit failure and cut their losses rather than waste even more money and time.

[54] Hock, *One From Many*, p. 242.
[55] Totten interview.
[56] Russell correspondence. Confirmed by Totten.

6.2.3 Success in Failure

Although admitting failure was difficult for Hock at the time, he was able to use the termination of BASE III to achieve a more important goal he had been pursuing for some time. When NBI was formed, the largest members had forced Hock to agree to a maximum cap on their quarterly service fees. These fees were based on the sales volume each member generated during the quarter, and were a key source of operating revenue for NBI. The larger issuers had already reached the cap, but NBI was still required to handle their ever-growing volume. Hock wanted to eliminate the cap, but since it was written into the bylaws, removing it required approval by a super-majority of the Board.[57]

Hock went to the Board with a plan about how to terminate BASE III, recoup the losses it had caused, and generally streamline NBI's operations. The plan had four points: The third version of BASE III would be abandoned, the entire project cancelled, and the advance license fees refunded; NBI would layoff a third of its "non-essential" employees; member fees would be reduced overall; but the cap would be removed and replaced with a sliding fee to be phased in over five years. Hock persuaded enough of the Board members to adopt the plan, thereby removing the cap, but the victory was bittersweet. On 22 January 1976, thereafter referred to as "Black Thursday," a third of NBI's non-essential employees were laid off.[58]

6.2.4 Returning to Purpose and Principles

In retrospect, both Hock and Russell admit that the central mistake of the entire BASE III project was actually a loss of focus. Hock wrote:

> The lesson finally sank in. We had never gone back to our purpose and principles to ask what card processing software for bank use had to do with "creating the world's premier system for the exchange of value." It had nothing to do with our purpose, or our belief about decentralization of function in pursuit of it. It was not a mistake of the people, it was a mistake of the leadership. We were extremely good at some things, we were very good at many things, but we were not good at everything. One of the principle arts of leadership is to make such distinctions and we failed to do so.[59]

Russell concurred:

> . . . we tried to become all things to all people, instead of doing that for which we were created. Writing and selling file management software is a completely different bag from managing and operating online payment systems.[60]

[57] Cleveland, pp. 11–12.
[58] Ibid.
[59] Hock, *One From Many*, p. 242.
[60] Russell correspondence.

Both Hock and Russell (as well as the rest of the early NBI leadership) believe that the primary purpose of a networked payment system is to provide a framework in which competitors can cooperate just enough in order to offer a service that no single member could have realistically offered alone. BASE I and II both supported that purpose, but BASE III did not. BASE I and II were simply coordination mechanisms, and neither constrained the members from adding proprietary features to their local systems in order to gain a competitive advantage. Although BASE III would have made it slightly easier for some members to coordinate with NBI's core systems, it also threatened to homogenize those members by supplying them with a single, common software program that would have been difficult or even impossible to augment with local innovations. Thus, BASE III upset the delicate balance between competition and cooperation that is so critical in a networked payment system.

More importantly, BASE III demonstrated that NBI had temporarily lost sight of what kind of system they were actually building. By developing BASE III, NBI was acting more like a computer systems vendor, a role for which they were neither created nor prepared. NBI was formed to build and operate an electronic payment system, not develop computer systems for their own sake. The latter was only a means to achieve the former.

NBI's leadership also learned a more general lesson from the BASE III project. Although there would be times when they needed to encourage the development of ancillary technologies to enhance the functionality of the network, they should constrain their input to only those mechanisms that enable *coordination* (e.g., data formats, communication protocols, and other standards). Construction of the actual technologies should then be left to firms that specialize in such devices, and they should enlist *multiple competing* firms so as to encourage further innovation. NBI applied this lesson well when they encouraged the development of inexpensive merchant dial terminals, the story of which will be discussed in Chapter 8.

6.3 Conclusion

In this chapter, we reviewed how NBI automated the clearing and settlement of interchange transactions, which was one of the key operational problems facing the BankAmericard system. We also discussed the BASE III project, which was an attempt to replace the punched-card accounting system developed by the BofA with a more modern mainframe computer program. This project ultimately failed, primarily because it was a significant departure from NBI's core purpose.

With the combination of BASE I and II, Hock was much closer to his dream of worldwide electronic value exchange. However, there were still a number of areas in which the system needed to expand. Hock wanted to enable cardholders to access any kind of funds they might possess, including deposits and investments, which implied that the system needed to expand outside the credit departments of the member banks. He wanted cardholders to have access to those funds anywhere in the world, which implied that the cooperative organization needed

to expand outside the US. It also implied a significant expansion of the computer systems and networks not only to handle the increased volume, but also to ensure a dependable service. These expansions, as well as others that were forced upon NBI, occurred throughout the 1970s, and it is to these that we now turn.

Chapter 7

Expanding the System: Organizational and Technical Changes

In the previous chapter, we reviewed the creation of NBI's BASE II, which automated the clearing and settlement of interchange transactions. With the combination of BASE I and II, NBI now had the necessary data processing infrastructure on which they could build Hock's worldwide electronic system for the exchange of value. However, to achieve this goal, the payment system needed to be expanded, both organizationally and technically.

In this chapter, we will examine the various ways in which the system was expanded throughout the rest of the 1970s, most of which were instigated by NBI, but some were also forced upon them. I will segment these expansions into two major sections: those involving the organization, and those involving the computer systems. However, in order to understand the explosive growth that occurred during the late 1970s, one must hold these two aspects together. The organizational expansions certainly triggered much of the growth, but if the computer systems had not also been expanded during this same period, the payment system would have quickly contracted again, or at least grown far more slowly.

7.1 Expanding the Organization

From 1972 to 1976, the NBI organization experienced three fundamental changes that would significantly shape the payment system's evolution. First, Hock and his staff replicated the NBI organizational structure at the international level, forming a new organization called IBANCO, of which the entire NBI system was a member. Second, NBI's prohibition on participating in both their system and Interbank's was repealed after a lengthly legal battle and Department of Justice review, creating the condition known as "duality." Third, both of these changes were woven together when the organization adopted the name "Visa," an action that became the catalyst for a rapid expansion of the membership, cardholder base, and sales volume. We will now examine each of these in turn.

7.1.1 IBANCO

From its inception, the BankAmericard licensing program was international, at least to some extent. Barclays Bank, Ltd., in the UK became the first international licensee in 1966, and by late 1972, banks in fifteen countries were participating in the program.[1] Although the BankAmericard Service Corporation (BASC, a wholly-owned subsidiary of BofA created to manage the card program) relinquished control of the domestic licensees to NBI in 1970, it still retained control over the international ones. Thus, the international organization remained roughly the same as the domestic one had been prior to NBI, and not surprisingly, experienced many of the same problems.

According to Hock, the international licensees attempted to break away from the BASC at the same time NBI was forming, but were unable to organize.[2] In 1972, after observing the initial successes of NBI, they made a second attempt, asking Hock to help create an international version of the organization. Hock was of course delighted, as this would further his ultimate agenda of creating a worldwide system for the exchange of value.

Determining the Structure

However, creating an international organization would be far more difficult than establishing NBI. Despite the differences between the US BankAmericard licensees, they still all spoke the same language, used the same currency, served a roughly similar culture, and participated in a common banking system. A global organization could assume none of those, and would thus need to separate the system elements that were truly universal from those that were specific to a particular region, country, or member. These universal elements were primarily Hock's set of principles discussed in Chapter 4, the rules governing international interchange, and the use of the worldwide marks.

Although every licensee had an economic interest in forming an international cooperative organization, they also feared that one country would amass too much political power, and thus control over the system. The chairman of the international organizing committee explained that,

> In structuring the new corporation, the committee sought to gain a balance between, on the one hand, equitable representation of all members and protection of the system from domination by a single member, country or region, and [on] the other, preservation of autonomy of each member's bank card program and each national bank card system, consistent with the interests of the worldwide system.[3]

This balance was obtained primarily by structuring the Board of Directors to ensure that no member, country or region had a majority of votes, but all had enough representation to

[1] Phillip Brooke, 'IBANCO, Ltd., Organized to Manage Worldwide BankAmericard Program', *American Banker* (20 September 1974), p. 1.

[2] Hock, *One From Many*, p. 196.

[3] Brooke, 'IBANCO Organized', p. 8.

block policy that was not in their best interest. Even the United States, which generated the vast majority of the system's volume, held a minority of seats on the Board. Unlike the NBI Board, the international licensees had less interest in controlling the daily operational details, leaving those decisions to the NBI staff.[4]

Studium ad Prosperandum, Voluntas in Conveniendum

After two years of negotiation, the structure of the new organization was established and nearly all the issues had been resolved. However, there were some critical disagreements still, primarily centering around Hock's principle that any qualifying bank must be admitted to the system. The international licensees had enjoyed exclusive rights to their country or region for many years, and the prospect of allowing their competitors to become members of a system they had spent money and time developing was unacceptable to some.

Hock scheduled a final meeting in San Francisco where these last issues would be discussed, but he expected that little would change. To achieve consensus, Hock realized he would need a device to give the licensees some perspective.

Prior to the meeting, Hock had a set of custom-designed cufflinks made for each of the representatives. As expected, the meeting itself achieved little, and the effort appeared to be a lost cause. As a final tribute, Hock took the licensees to a restaurant in Sausalito, the same town where he had developed his principles for the NBI organization four years earlier. After dinner, he placed the cufflinks before each representative and began one of the most important speeches of his career. His recollection of it is so powerful, it is worth quoting in full:

> . . . a small gift was placed before them as I concluded. "It is no failure to fall short of realizing such a dream. From the beginning it was apparent that forming such a complex, global organization was unlikely. We now know it is impossible, notwithstanding two years of exceptional effort. Not knowing with certainty how today's meeting might end, we felt compelled to do something that would be appropriate no matter what happened. Would you please open the small gift on the table before you?" As they each opened a small, beautifully wrapped box and began to examine the contents, I quietly continued.
>
> "We wanted to give you something that you could keep for the remainder of your life, as a reminder of this day. On one cufflink is half of the world surrounded by the Latin phrase, 'Studium ad prosperandum'—the will to succeed. On the second cufflink is the other half of the world surrounded with 'Voluntas in conveniendum'—the grace to compromise. We meet tomorrow for the final time to disband the effort after an arduous two years. There is no possibility of agreement. As organizing agent, we have one last request. Will you please bring your cufflinks to the meeting in the morning? When it ends, each of us will take them with us as a reminder for the remainder of our lives that the world can never be united through us because we lacked the will to succeed and the grace to compromise. But if, by some miracle, our differences dissolve before morning, this gift

[4] Russell interview.

> will remind us to the day we die that the world *was* united because we *had* the will to succeed and the grace to compromise."
>
> There was a moment of profound silence as they examined their gift. It was shattered by one of my more exuberant Canadian friends, may his soul rest in peace, who rose with a huge grin and exploded, "You miserable bastard!"[5]

The next morning, the licensees reached consensus on every outstanding issue, and they unanimously agreed to form the new organization.

Although this story sounds almost too good to be true, the cufflinks and the basic points were verified by other sources.[6] However, it is uncertain what role the cufflinks actually played in encouraging compromise. Russell recalls that the licensees agreed to allow competitors to join only after it was decided that new members would pay higher fees to compensate the original licensees for their efforts in developing the system in their respective regions. Ultimately, the degree to which the cufflinks engendered cooperation cannot be known without interviewing the licensees in attendance, which is beyond the scope of this project.

Formation

In June 1974, IBANCO was organized as a for-profit, non-stock Delaware corporation, just like NBI.[7] NBI continued to manage the US members, but became a member of Ibanco, along with other national associations and individual licensee banks. Hock became president of Ibanco as well as NBI, and Ibanco contracted with the NBI staff for the management of Ibanco's operations. A new set of international bylaws and operating regulations were created from those used by NBI, but individual regions and countries were free to enforce their own, additional rules, provided they did not contradict the international ones.

The creation of Ibanco was an important expansion towards achieving Hock's goal of a worldwide electronic value exchange system, primarily because it created a framework of rules in which all BankAmericard participants could both cooperate and compete. Under the BASC, each licensee negotiated their own terms, and there were no mechanisms to enforce the few rules that existed. Under Ibanco, there were clear and consistent rules as to the interchange reimbursement fee, dispute resolution, the use of marks, design of cards, format of the drafts, as well as penalties for violating the rules. These rules were necessary to provide a cohesive, networked, worldwide payment service.[8]

[5] Hock, *One From Many*, pp. 210–211.

[6] The cufflinks themselves are pictured in Chutkow, p. 138. Russell confirmed the dinner and the presentation.

[7] Brooke, 'IBANCO Organized', p. 1. IBANCO reportedly stood for "International Bankcard Company," but was commonly referred to only by the shortened name, and often in title case after first use (i.e., Ibanco).

[8] Of course, rules are always open for interpretation and dispute, and Visa's mechanisms for resolving those disputes will be discussed in Chapter 10.

7.1.2 Antitrust and Dual Membership

As the organization expanded internationally, it also expanded domestically, though not entirely by choice. Throughout the early 1970s, NBI was embroiled in an antitrust lawsuit brought by a bank that was both a member of NBI and Interbank. The case posed a question regarding this new type of organization that had no obvious answer: should banks be allowed to join multiple bankcard systems? Since the member banks actually own and govern these systems, would this lead to more or less competition?[9]

Historical Context

The answers to these questions were further complicated by a division of the membership originally created by the BASC licensing structure. Recall that when BofA franchised the program in 1966, it gave most banks exclusive licenses for their territory. These banks saw the exclusive right to issue the cards as a great benefit, as they typically considered the card program not as a profitable business venture of its own, but as a method for attracting new customers and thus deposits. However, most banks, especially those in states with restrictive branching laws, often needed the help of other, smaller banks to enlist a sufficient number of merchants to make the system attractive and viable. Thus, the licensee banks contracted with these *agent* banks to sign up additional merchants and process their transactions. Many licensee banks also allowed non-competing agents to "re-issue" their cards to the agent's depositors, but the licensee held the receivables and thus enjoyed all the interest revenues.

When NBI formed, it maintained this distinction, labeling the issuing banks *Class A* members and the agents *Class B*. The two-hundred-plus Class A members were full owner/members of NBI, while the thousands of Class B members were merely "sponsored" by their respective Class A member.[10] By the late 1970s this distinction had eroded and was formally eliminated, but during the first part of that decade, it was fundamental to understanding the issue of duality.

By the time NBI formed, a small number of Class A banks had also joined Interbank, primarily to hedge their bets on Hock's tenuous new organization. NBI immediately put a moratorium on this practice, fearing that Interbank would learn of NBI's plans through the dual members. One of the dual Class A members, Worthen Bank and Trust of Little Rock, Arkansas, warned NBI's management that they would challenge any attempt to make the ban permanent. However, Hock was adamant, and in October 1971, NBI adopted a new section into their bylaws, number 2.16, which prohibited any member that issued a BankAmericard (i.e., Class A member) from participating in any other bankcard association.[11] As promised,

[9] Information on duality and the antitrust case comes from the following sources: Hock, *One From Many*, pp. 181–194; C Frederic Wiegold, 'NBI Appeals to Court for Chance to Prove Dual Membership Ban is Not Group Boycott', *American Banker* (2 January 1973), p. 1 and other trade press articles; interviews with Katz, Honey, Russell, Powar, Conway, and Derman.

[10] By 1973, there were 250 Class A and 4,410 Class B members (various articles in *American Banker*).

[11] Although Hock was adamant, the rest of his staff was actually divided on the issue, and there was considerable internal debate. See Hock, *One From Many*, pp. 182–183. Also confirmed in Katz and Russell interviews.

Worthen quickly filed a private antitrust suit, claiming that the ban was a horizontal restraint on trade deserving an immediate injunction.[12]

It is important to understand that although Worthen wanted to be a member of both systems, it was not as interested in *issuing* both types of cards as it was in *acquiring* both types of transactions. Worthen's primary complaint was that NBI's ban applied to Class A members only; Class B members were still allowed to acquire transactions for both systems. Because merchants wanted to deposit all their various bankcard transactions with the same institution, Worthen argued that Class B members had an unfair advantage in the acquiring business.

The Arguments

Whether the ban on dual membership would result in more or less competition was complex to answer because it depended on one's perspective. From the perspective of an individual bank, the ban seemed to restrict competition, as Class A members had less freedom than the Class B members. However, from the perspective of the entire payment system, the ban seemed to increase the potential for competition, as the two systems would need to offer competitive fees and develop innovations to attract and retain their members. If banks could be members of both systems, there would be less incentive for them to encourage competition at the system level, as they would be owners and tax-payers of both. Furthermore, Hock argued that there would eventually be enormous pressure to merge the two systems, and little incentive for banks to break away to form a new competitor. Thus, whether the ban encouraged competition depended largely on whether one wanted to optimize for competition between the individual banks, or between the systems themselves.

Hock was focused on the system as a whole, and as such, was strongly against duality. Furthermore, he also opposed the idea of any single agency or organization dictating what was in the best interest of the public. In a 1974 speech, he observed that,

> The proponents of several views seem to make an assumption that such matters as the public interest and social cost can best be judged by whatever entity they deem suited to the task ahead. The real question is, who is best suited to judge such matters? Congress? The Executive Branch of Government? The Federal Reserve? Commercial Banks? Savings Banks? Bank Card Organizations? The Consumers Union? The Justice Department? Or is it the public? And if the public, how can it have any opportunity to exercise its judgement except by the only effective method which has ever been found; that is, by choosing freely among a variety of competing services with complete information about the costs, practices and benefits of each. My strong conviction is that the public should choose and my great fear is that they may never have the opportunity. If solutions are forced upon them it matters little which organization does so, for ultimate abuse of such power will be inevitable.[13]

[12] Hock, *One From Many*, pp. 181–184.

[13] Hock, 'Electronic Funds Transfer or Electronic Value Exchange?', p. 3.

Although his claims are certainly open to critique, this quote allows us to peer into Hock's way of thinking. One of the most salient features of natural systems for Hock is how they function so effectively without centralized control. Here he expresses the common libertarian belief that a free market is far more effective than myopic regulation, which is, for a libertarian, the most pernicious form of centralized control. Hock wanted to see many competing payment networks, and he feared that without a ban on dual membership, no new payment system would ever arise. Furthermore, NBI and Interbank would eventually be pressured to merge, and the resulting monopoly would ultimately abuse its power. In the end, consumer choice would be restricted, not enhanced.

Trials

On 19 July 1972, the district judge sided with Worthen, declaring the ban on dual membership to be a "group boycott" and thus a horizontal restraint of trade that was *per se* illegal under antitrust law.[14] By deeming it "*per se* illegal," the judge considered it to be so obviously anticompetitive that no detailed economic or business review of the system was necessary. The judge also granted an immediate injunction prohibiting NBI from enforcing the ban, but delayed the award of damages until after the appeals process.

NBI immediately appealed the case, arguing that because this type of venture was so new and untested, the ban should be evaluated under the "rule of reason," meaning that the court should conduct a full business and economic analysis of the system in order to determine if the ban was indeed anticompetitive. In a friend of the court brief, the Department of Justice (DOJ) also encouraged a full trial to determine the merits of the policy.[15] In September 1973, the appellate court agreed, sending the case back to the district court for a full trial under the rule of reason.[16] Worthen appealed to the Supreme Court to uphold the *per se* ruling and injunction, but was denied a hearing.[17]

The costs of an extensive trial were unappealing to both sides, so NBI and Worthen eventually reached a settlement. NBI agreed to extend the dual membership ban to Class B banks as well as Class A, but wait to enforce it until they received a "business review" letter from the DOJ, stating their opinion and enforcement intentions.[18] Thus far, NBI's arguments had met with favorable responses from the DOJ, but the DOJ's head lawyer had recently stepped down, and the new leadership was far more skeptical. After a lengthy investigation, the DOJ

[14] Wiegold, 'NBI Appeals to Court'.

[15] C Frederic Wiegold, 'Worthen Suit Against NBI's Ban Needs Full Trial', *American Banker* (12 July 1973), p. 1. For Worthen's response, see 'Worthen Opposes Call for Full Trial', *American Banker* (23 July 1973), p. 1.

[16] 'Court Upsets Ruling Against NBI Dual Card Ban', *American Banker* (25 September 1973), p. 1.

[17] 'Plea for New Hearing by Worthen is Denied', *American Banker* (25 October 1973), p. 1; 'Worthen Asks High Court to Review Reversal', *American Banker* (11 December 1973), p. 1; 'High Court Refuses Worthen-NBI Review', *American Banker* (20 Feb 1974), p. 1.

[18] Hock, *One From Many*, p. 190. See also 'NBI Maps New Rules Banning Dual Membership', *American Banker* (17 September 1974), p. 1; Phillip Brooke, 'NBI Moves to Reinstate Contested Ban of Dual Membership', *American Banker* (20 November 1974), p. 1.

issued a letter on 7 October 1975 that was so carefully hedged "as to leave unresolved the legal permissibility of an effective prohibition against dual membership."[19] They neither explicitly condemned nor condoned the ban, leaving NBI a target for further antitrust litigation.

Duality

Although NBI felt they could defend the ban in court, the costs of losing were enormous. Under US antitrust law, a convicted corporation was liable for treble damages, and its officers could face jail sentences. Hock had little chance of convincing the Board to enforce the ban, and as we shall see, eliminating it could help achieve another, perhaps more important goal. Thus, the bylaw was summarily removed in June 1976, ushering in a state known as "duality" in the US.[20] Banks were now free to join both NBI and Interbank, and most did. However, this resulted in a larger and more important expansion for NBI than Interbank, as the latter's membership was not only significantly larger than the former's, it also comprised the largest and most important banks in any given area.

The subsequent history of bankcard systems has shown that Hock's predictions regarding the effects of duality were entirely correct. Most banks went dual within six months, no competing system arose, and dual members found little incentive to pay for innovations in one system that they would only have to pay for again in the other. Reflecting on it in 1979, Hock told an interviewer:

> It won't be until 1981 when banks can sit back and ask, 'What hath God wrought?' or more properly 'What hath the Department of Justice wrought?' But already you can see some changes. When banks were either Master Charge *or* Visa issuers, every time one card association came up with an innovation, the members of the other would press their association to come up with a competitive answer. Now, rather than press the other association for innovation, some banks just criticize whichever is the innovator. As joint owners they can now solve competitive problems by inhibiting the innovator rather than by spurring the laggard. I have seen a lot more of this in the last year than I have in the past and it disturbs me.[21]

The ultimate irony came in 1998, when the DOJ sued both Visa USA and MasterCard, claiming that their common governance was creating anticompetitive effects. Additionally, the DOJ sought to eliminate rules that restricted members from issuing American Express, or any other proprietary card. The court sided with the DOJ on the latter issue, but not the former. In any case, it was too late to separate the two organizations.[22] Most recently, MasterCard has reorganized into a publicly traded corporation, and Visa has announced similar plans. These

[19] Baxter, p. 587.

[20] 'A Second Card In Your Future?', *ABA Banking Journal* (August 1976), p. 54. Other countries such as Spain and France did not allow such duality. The effects this had on the system evolution and competitive landscape in those countries would be an interesting focus for further research.

[21] Streeter, p. 70.

[22] Evans and Schmalensee, pp. 281–284.

moves could be viewed as attempts to evade further antitrust litigation. However, as long as there is common governance, whether by irrevocable rights of participation or stock ownership, the possibility of further litigation remains.

7.1.3 Adopting the Name VISA

The final organizational expansion to discuss was the adoption of the name we know the system by today—VISA.[23] It might seem strange to refer to a name change as an "expansion," but the adoption of the Visa name not only ties together the two structural expansions already discussed (Ibanco and duality), it also became the catalyst for an explosive growth in the membership, cardholder base, merchant acceptance, and sales volume, all of which would quickly make Visa the dominant bankcard system in the world.

The Need for a New Name

When the BofA originally licensed the BankAmericard program, they insisted that the name and the blue-white-and-gold bands design (the BWG mark) be used on every card, regardless of issuer. This was necessary to build the universal recognition and acceptance needed by any payment system. However, the licensee banks always resented having to use the name "BankAmericard," as it evoked the name of the largest and most powerful bank in the nation, and made their banks look subordinate to it. Although the BASC eventually allowed the licensees to add their own name to the card, the BASC still required the BankAmericard name to be in larger, more prominent type.

If the BankAmericard name was merely contentious for the domestic licensees, it was absolutely unacceptable for most of the international ones, as it evoked not only the name of a powerful multi-national bank, but also a powerful foreign country. Thus, most of the international licensees negotiated the use of an entirely different name on the card, though they retained the same BWG mark. As a result, the BankAmericard was known as "Barclaycard" in the UK, "Sumitomocard" in Japan, "Chargex" in Canada, and "Carte Bleue" in France.[24]

This confusing array of names inhibited the universal acceptance that Hock desired. Merchants who saw few international travelers were unsure if they were allowed to accept cards that looked similar but had different names. If Hock was to build the truly universal, worldwide value exchange system he envisioned, he needed to create a new name that would be acceptable and appealing to cardholders, merchants, and the entire Ibanco membership.

However, there was another, perhaps more important reason to change the name. When Hock was forming Ibanco, he and NBI's general counsel, Bennett Katz, approached the BofA

[23] As we shall see, VISA is technically an acronym, but it is commonly written in title case after first use.

[24] Some international licensees followed the domestic model of merely adding the licensee bank's name (e.g., "Banco de Bilbao BankAmericard"). For a pictorial catalog of the various international cards, see "Visa' will replace 'BankAmericard' Worldwide', *The Nilson Report* 147, p. 2. Also note that the domestic Carte Bleue card was a national debit card only, so they issued a separate BWG card for use outside France.

about transferring ownership of the BWG mark to the new organization. The BofA responded that although they would be happy to license the mark to Ibanco, transferring ownership would not be possible, as the mark was too closely associated with the name BankAmericard, which was in turn closely associated with the Bank of America. Hock and Katz countered by arguing that if Ibanco could associate a different name with the mark, BofA should then have no objections to transferring ownership. BofA concurred, and signed an agreement to relinquish ownership of the mark to Ibanco, without additional compensation, provided that ninety-five percent of the membership, comprising ninety-five percent of the volume, adopted a new name within five years (i.e., by 1979).[25] Note that at the time, the BofA generated more than five percent of the total volume, so this agreement gave the BofA the right to block a new name and card design if it was contrary to their interests.

Although the reasons for a new name are generally acknowledged by all sources, the creation of that name is a matter of contention. According to Hock, he and his staff developed a set of principles the new name needed to embrace, and then held an internal, "self-organizing" naming contest. The name "Visa" was suggested many times, but because there were dozens of different recollections as to who originally suggested it, everyone was given the credit.[26] Honey contends that he suggested the name to Hock for the original debit card, but Hock reserved it for the overall name of the product and organization.[27] The Hoefer, Dietrich, and Brown advertising agency of San Francisco sued NBI/Visa in 1977, claiming that they suggested the name during an advertising pitch, but settled out of court when documents produced during discovery showed that the name had been generated internally prior to their presentation.[28]

Redesigning the Card

Regardless of how the name was generated, NBI faced the arduous task of convincing the membership to adopt it. The domestic members would generally welcome it, but convincing BofA and the international members would be particularly difficult, as those banks currently had names on their cards that were strongly identified with their own organizations or services. To accommodate their concerns, Hock proposed what was known as the "tri-level marks" concept, described here in a speech from 1976:

> Card designs are... self-limiting. They do not accommodate evolving services or equitably balance the legitimate needs of involved parties. The ideal design would accommodate marks owned by an individual bank, representing services available

[25] Katz and Honey interviews. See also Hock, *One From Many*, p. 216.

[26] Ibid., pp. 217–219. Goldsmith claims that he was the one who originally suggested the name to Hock (Goldsmith interview).

[27] Honey interview.

[28] 'Visa Concept Stolen By NBI Says Creative Agency', *The Nilson Report* 163, p. 1; 'Visa Settles Ad Agency Suit Over Visa', *The Nilson Report* 208, p. 2. NBI actually compensated a number of financial institutions that had already used the Visa name in one way or the other. See Chutkow, p. 217; and 'BankAmericard's Conversion to Visa Will Cost $9.2 Million', *The Nilson Report* 160, p. 1.

> only from it. It would accommodate marks owned by an association of banks within a country, representing services good only in that country. It would accommodate marks owned by an international organization representing services which are universal throughout the world.[29]

With this new concept, Hock made an important philosophical shift regarding the nature of the card. He now saw it as a substrate that could carry any number of marks identifying not only the issuer, but also a range of services available to the cardholder. It was still an identity device, but it would now identify the cardholder to any of the services advertised upon it.

The task of redesigning the card and implementing the name change fell to a product developer named Tom Honey. While pursuing a graduate degree, Honey had been fascinated by an article on the future of electronic funds transfer (EFT) written by John Fisher of City National Bank and Trust (Columbus, OH), the same bank that conducted the full EFT test discussed in Chapter 5. Honey went to work for Wells Fargo Bank and the Bank of California, installing one of the first ATMs in the US in late 1971, and becoming generally acquainted with all the new banking technologies. He eventually interviewed with Hock, who saw in Honey a kindred passion and vision. Honey joined NBI in 1973 and developed their first debit card, named Entrée, which will be discussed in detail in Chapter 9.

In keeping with the tri-level marks concept, Honey and his team developed a new design for the card and merchant signs:

> We kicked around various design options until it became clear: we would put the VISA name in the middle white band of both the card and merchant decal and allow every issuer to use the top blue band of the card for their own name and service indicia. That solved both our international issuer problem and the Bank of America problem. We just added VISA to the card name.[30]

This solution not only placated BofA and the international members, it also solved a nagging problem that had existed from the early days of the licensing program. For reasons unknown, merchant signs had always featured the BankAmericard name in the white band, rather than in the blue band as it was on the card. With the new design, the merchant signs and cards were now identical.

More importantly, Honey's new design also included a phased reduction in the size of the marks so that the issuing organization could control more of the front of the card. The first phase would simply add the Visa name to the white band, but the second phase would reduce the colored bands and Visa name to a small rectangle either in the center or the lower-right corner of the card.[31] In a proposed third phase, which has not yet occurred, the entire mark would be moved to the back of the card where the various ATM network marks appear today.[32]

[29] Dee Hock, 'Bank Cards: Today's Popular Song or An Unfinished Symphony?' (14 September 1976), p. 9.

[30] Honey correspondence.

[31] It was first moved to the center, and then to the lower-right where it remains today, though Visa has recently removed the colored bands from the mark. Pictures of the centered placement can be seen in Chutkow, p. 215.

[32] This was one of the key reasons cited by Citibank for leaving the Visa system in 1998. See Evans and Schmalensee, p. 171.

With these phases, Honey was making a subtle but important distinction between *brand* and *mark*, asserting that Visa was actually the latter and not the former. The issuing organization established the brand, and Visa was simply a mark that was added to identify an available service. This allowed banks to issue locally-branded cards that could access a number of different financial service networks. In many ways, the Visa mark is similar to other so-called "ingredient brands," such as Intel Inside®, or Dolby® sound.[33]

Furthermore, the tri-level marks concept also allowed for a compromise that would eventually entice most of America's major retailers to accept the card. Recall that the large-scale merchants had always resisted bankcards because they threatened the loyalty gained through the merchant's proprietary card. However, since those merchant-specific cards were rarely profitable, a bankcard that allowed merchants to sell off their receivables to a bank, while retaining their identity on the card, would be an acceptable compromise. The new card design allowed for just that; Visa members could offer merchants a card with the merchant's name in the blue band, and enable any set of incentives the merchant wanted to offer customers. Once the Visa marks were reduced in size, the merchant could further customize the look of the card, making it appear like they simply added the Visa marks to their original, proprietary card.[34]

Implementing the Name

Redesigning the card and merchant signs was relatively easy, but getting the members to implement these changes would be much more difficult. Every merchant acquirer had to replace every sign and decal at each of their merchant locations, and in Japan this included thousands of expensive, lighted Sumitomocard signs. Every issuer had to print and distribute new cards to every cardholder. All the members had to replace all their paper forms, and run advertisements informing the public that the name was changing. All of this implied a significant cost that created a natural incentive for the members to resist a rapid changeover. NBI's management realized that they needed a catalyst that would push at least the domestic members to implement the change quickly.

Although I have presented the duality issue and the name change in separate sections, they were actually occurring at the same time and must be understood in relation. When the DOJ issued the review letter, NBI's management knew that they must remove the ban on dual membership. However, they also saw the opportunity to create the catalyst they needed for a rapid implementation of the Visa name, and set about mixing the ingredients.

In May 1976, the NBI Board approved the name change to Visa, and the Ibanco Board followed suit in August. In June 1976, NBI lifted the ban on dual membership and approached a few strategic Interbank members, inviting them to join NBI. In September 1976, NBI formally

[33]Honey interview.

[34]For a report on the card done for the Woolworth Company, then the fourth largest retailer in America, see 'Woolco Issues own Visa Card', *The Nilson Report* 266, p. 1.

announced the new Visa name.[35] In March 1977, Ibanco was renamed to Visa International Services Association (creating the recursive acronym VISA), and NBI was renamed to Visa USA. The membership was then given six months to replace their merchant signs, and an ample two years after that to reissue the new cards. Visa USA then began a multi-million dollar national advertising campaign, supported by additional local member advertising, informing the public that BankAmericard was now becoming Visa.

Those few strategic Interbank members, now Visa USA members as well, then began sending out applications to consumers in their area, including those holding BankAmericards issued by their local competitors, offering to convert consumers to this new Visa card they were hearing so much about. One of those new members was Citibank, then the second largest bank in the nation, and they went one step further by sending out applications across the *entire nation* starting in the fall of 1977.[36] Consumers, who were being barraged with ads promoting the new name, were expecting something related to the new card, and typically did not notice that the offer came from a different bank, nor did the applications make that particularly obvious.[37] The old BankAmericard issuers, who were planning to phase in their new Visa cards slowly over two years, saw what was happening, realized they could lose their cardholder base, and rushed to re-issue the new Visa cards. Meanwhile, they had little time to apply for membership in Interbank, much less think about issuing Master Charge cards.[38]

The strategy was highly effective, simultaneously forcing a rapid implementation of the new name, and delaying the Visa USA members from joining Interbank and issuing Master Charge cards. Soon afterward, a consumer opinion survey found that the sample not only understood that BankAmericard had become Visa, but a large portion of respondents also thought that *Master Charge* had become Visa as well.[39]

The creation of Ibanco, the establishment of duality, and the name change to Visa all combined to create an explosive growth that quickly made Visa the dominant card system in the world. By the first quarter of 1978, Visa had surpassed Interbank's sales volume both in the US and worldwide.[40]

[35] "Visa' will replace 'BankAmericard' Worldwide', *The Nilson Report* 147, p. 2.

[36] 'A Visa-Card Offensive Angers the Opposition', *Business Week* (5 September 1977), p. 31. See also 'Citibank Goes BankAmericard', *The Nilson Report* 157, p. 2; and for Nilson's accusation that Citibank was sending out applications indiscriminately, see 'Citibank Mails Bank Card Invitations to Phone Book Listings', *The Nilson Report* 180, p. 1.

[37] *Business Week* reported that "an estimated 60% of the 55-million holders of national bank cards do not know which bank issued their cards or processes their accounts." See 'Bank Cards Take Over the Country', *Business Week* (4 August 1975), p. 53; and 'Visa-Card Offensive Angers'.

[38] Although Nocera mentions the concurrency of duality and the name change, he presents it merely as context for the Citibank national issuance, and not as a mechanism used by Visa to accomplish a rapid implementation of the new name. See Nocera, pp. 144–148. Chutkow notes that the Citibank issuance helped to increase the speed of the conversion, but does not attribute any role to Visa, nor does he discuss the specific changes made to the computer systems during this time that enabled Visa to handle the growth. See Chutkow, p. 219.

[39] Honey interview.

[40] 'Visa's Climb to No. 1 in Cards', *Business Week* (17 July 1978), p. 26.

7.2 Expanding the Computer Systems

As the organization expanded, so did the computer systems, and these two developments must be understood in relation. Although the organizational changes just discussed contributed to Visa's explosive growth during the latter 1970s, the critical changes made to BASE I and II during this same period allowed them to *handle* that growth, and ultimately led to even further expansion of the payment system.

These changes can be grouped into three principal areas. First, BASE I's capacity was significantly increased by re-implementing it on IBM hardware and the Airline Control Program (ACP) originally developed for the Sabre airline reservation system. Second, another data center was opened on the East Coast, running in parallel with the one in San Mateo. Third, the computer systems were extended internationally to provide automated authorizations, and eventually clearing and settlement, to all Visa International members. We will now examine each of these in turn.

7.2.1 Expanding BASE I's Capacity

The dependability of BASE I has always been a top priority for Visa's computer systems staff. Visa's existence largely depends on their ability to offer a payment service upon which cardholders, merchants, and members can rely, and the authorization service is the primary and most common point of contact for those groups. If BASE I is unavailable, or misbehaves in some way, customer confidence can quickly erode, devaluing the card as a payment device. Visa employees call this the "back of the wallet" effect—if a cardholder is improperly denied an authorization, the cardholder typically moves that card to the back of the wallet, and will choose other methods of payment for future transactions.[41]

Capacity in Real-Time Systems

As discussed in Chapter 5, the design of the original BASE I included a few key mechanisms to ensure the availability of the authorization service. All critical telecommunications and computer hardware had redundant backups. The central switch also had the ability to stand-in and approve transactions when the issuer's host computer or personnel did not respond. NBI could also stand-in for the acquirers, answering authorization calls after the acquiring centers closed for the night.

However, there was another threat to availability against which redundant hardware and logic could not guard—overwhelming the system's *capacity*. In any real-time, transaction-processing system, there is a somewhat fixed number of transactions the system can process at any given time. Once the concurrent transaction load reaches this amount, the system cannot accept more, and becomes unavailable until the existing transactions are processed. Often a

[41] Elliott interview.

queueing mechanism is used to hold new transactions until they can be processed, but eventually the CPU spends so much of its time switching between tasks and managing the queue that it simply cannot accept more transactions. If the software developers anticipate this condition, the system will gracefully reject new transactions, asking the sender to try again later; if they do not, the system often crashes.

For BASE I in particular, the floor limits provided a crude but effective mechanism for tempering the transaction load. Each merchant had a limit under which the transaction was automatically authorized. NBI also established interchange floor limits for the acquirers, under which the acquirer could authorize the transaction without contacting the issuer. However, the interchange floor limits were removed shortly after BASE I was put into production, and the merchant floor limits were reduced in an effort to stem fraud. Thus, the transaction load began to increase.

However, it is important to emphasize that the main threat to capacity is not the aggregate transaction load experienced over a given period, but the peak. Authorizations in particular have a more or less predictable ebb and flow cycle, with dramatic peaks occurring at certain times, and on certain days. However, how high those peaks will reach can only be estimated, and if the estimates are too low, the system will become unavailable at the most inopportune time.

BASE I was just over a year old when the system capacity was exceeded by an unexpectedly large peak:

> ...the day after Thanksgiving [1974] was the biggest shopping day of the year. Now over time that's changed a little, but for many years, the Friday after Thanksgiving was *the* volume day. Well anyway, BASE I died a horrible death; several times...they would have an [acquirer] that would choke them to death and cause it to crash. Well instead of shutting down the big users, bringing it back up and letting it start to breathe again, they just brought it all back up and guess what? It went right back down again. That day was one of the most horrible in Visa.[42]

An experienced systems developer named Frank Fojtik led a team to expand BASE I's capacity. Fojtik was exactly the sort of technical person Hock admired. Speaking with a no-nonsense Texan drawl, Fojtik remarked that his technical expertise was the result of "a whole lot of scars I got on my ass" while working with cutting-edge technology, and he summarized the entire Visa mission by saying, "if you can't recognize it, it ain't worth much; if it don't work, it's worth even less." Fojtik had been recently hired from Singer, where he had worked on electronic cash registers, and developed a number of networked, transaction-processing computer systems.

[42] Fojtik interview.

DEC or IBM?

Fojtik and his team were able to reorganize the structure of BASE I to achieve enough capacity to survive the 1975 Christmas shopping season, but they all knew it was merely a temporary solution. After the ban on dual membership was lifted in 1976, both the aggregate and peak volumes dramatically increased, and the systems group was now faced with a difficult decision: should they upgrade to the new DEC PDP-11/70; or should they switch over to the IBM System/370 mainframe line? Hock, in his usual style, declared that the decision would be made by all the key technical staff after considering the arguments presented from both sides. However, instead of letting the IBM and DEC sales people in the room, he assigned members of his own technical staff to make the case for each alternative, arguing that his own staff would think in terms of what was best for Visa, and would ultimately need to believe in the solution. Fojtik served as DEC's champion, and Derman pitched IBM.[43]

The argument for upgrading to the next PDP-11 was relatively easy to make, as it featured the same type of architecture and thus required little change to the BASE I software. However, DEC was suspiciously vague about when another, more powerful PDP-11 might be available, leading the technical staff to wonder about the long-term prospects of that platform.

The argument for moving to the IBM System/370 line rested primarily on their seamless upgradability. If their current mainframe ran out of capacity, it could be replaced with another more powerful one without any changes to the application software. IBM had a number of 370s already in production, and more planned for the future. More importantly, IBM had become the standard in the banking world, and would thus be easier to sell to the Board. But this approach also required a re-implementation of the BASE I system, as the IBM environment was completely different.

Hock held a lengthy meeting, listening to the arguments, pitting his staff against one another, and challenging all their assumptions. However, it was mostly a charade—Hock had already made up his mind earlier, but wanted his staff to come to the same conclusion. Chuck Russell conveyed how the decision was actually made:

> Dee was leaning toward re-upping with DEC, but before so doing, decided to give IBM one more shot. He asked the SFO [San Francisco] IBM chief, Bob Irwin, to come in with his right hand man, Roger Peirce, and meet with him and me to tell us why we should go with IBM.
>
> Now you have to understand that, at that time, SFO was considered *Siberia* by IBM... Bob no longer "fit in" at IBM because he was considered "uncouth" by the then button-down culture, which prevailed during the '70's and '80's at IBM... [So] Bob was banished to SFO... [but Bob and Roger] were/are two of the brightest individuals I've ever worked with.

[43] Information on the DEC vs. IBM decision comes from interviews with Derman, Fojtik, Peirce, Schramm, and Russell. There is some disagreement as to who voted for which alternative, but as we shall see, it mattered little in the end.

> Get this picture. Bob and Roger walk into Dee's office. Bob is reasonably well dressed, but he's wearing well-polished *GI paratrooper boots*! Roger has a button-down shirt on, but it looks like he slept in it, and his hair apparently hadn't seen a comb in weeks! This duet put on a presentation right out of Tom [Watson] Sr.'s book. Blunt, honest, brief and to the point, and not one ounce of BS or political correctness. It would have curdled the button-down set's milk!
>
> At the end of their presentation, Dee, in his usual fashion, commenced hassling them. After two or three minutes, Bob got up and said, "Look Dee, if you're too f—— dumb to understand the difference between what IBM brings to the table, vs. what DEC can do, you're wasting my time," whereupon both he and Roger walked out of Dee's office and slammed the door behind themselves.
>
> Dee, rarely at a loss for words, was speechless for a few moments, then he turned to me, and said, "Hell, chief, if they feel that strongly, they are the people to go with."
>
> And that's how the decision was made.[44]

Airline Control Program

Thus it was decided to re-implement BASE I on the IBM hardware platform, but this resulted in a new dilemma: which operating system should they use? IBM offered a few choices at this time, the most common being MVS, but this was designed primarily for controlling batch-oriented applications, not real-time systems. In 1969, IBM had released what became their most well-known transaction processing monitor, called the Customer Information Control System (CICS).[45] However, Peirce thought that it was not reliable enough at the time for Visa's needs. It was also relatively inefficient, and would have required an excess of powerful and expensive hardware to achieve the necessary capacity.[46]

However, there was another, less well-known operating system offered by IBM that Peirce thought would be perfect: the Airline Control Program (ACP). The design of ACP was originally developed for the Sabre airline reservation system, which was a joint project between IBM and American Airlines in the early 1960s.[47] After Sabre went into full production, IBM

[44] Russell correspondence.

[45] This was actually one of IBM's first software "products" sold independently of the hardware. See Paul E Ceruzzi, *A History of Modern Computing* (Cambridge, MA: MIT Press, 1998), p. 106. Note that CICS is actually a subsystem designed to run on MVS, not a full operating system itself.

[46] Peirce interview.

[47] For information on Sabre, see Duncan G Copeland, Richard O Mason and James L McKenney, 'Sabre: The Development of Information-Based Competence and Execution of Information-Based Competition', *IEEE Annal of the History of Computing* 17:3 (1995); John R Knight, 'A Case Study: Airlines Reservations Systems', *Proceedings of the IEEE* 60:11 (November 1972); Robert V Head, 'Getting Sabre Off the Ground', *IEEE Annal of the History of Computing* 24:4 (2002); and Campbell-Kelly, pp. 41–45. For a detailed review of ACP/TPF, see J E Siwiec, 'A High-Performance DB/DC System', *IBM Systems Journal* 16:2 (1977), pp. 169–195; and Thomas W Scrutchin, Jr, 'TPF: Performance, Capacity, Availability', in: *IEEE Compcon 1987 Digest of Papers* (Washington, DC: IEEE Computer Society, 1987), pp. 158–160. Additional information on ACP comes from interviews with Peirce, Fojtik, Totten, Reid, vonGillern, and Boston. Although the name "Sabre" is often written in all capitals, it was never an acronym, and was actually derived from the Buick LeSabre.

enhanced the system based on their initial experience, and repackaged it in a more generic form, known as Programmed Airline Reservation System (PARS). PARS was adopted by most of the major US airlines, and further enhanced based on its observed performance. The operating system that controlled the PARS application programs, which comprised most of the code base, was broken out and licensed separately as ACP beginning in 1972.[48] It was eventually renamed Transaction Processing Facility (TPF), and it is still the heart of most airline, hotel, and car rental reservation systems.

ACP is a rather unique operating system and database in one. It was specifically designed for the airline reservation environment, meaning that it is extremely adept at processing an unpredictably-large number of simultaneous transactions, each of which require little CPU time and limited but specialized operating system services. The goal of ACP was to enable the applications to process most transactions within just a few seconds, allowing the reservations clerk to keep up a normal flow of conversation with the customer.[49] In fact, the ACP designers considered the reservations clerk, as well as the system operators, to be crucial elements of an overall system comprised of both humans and machines.[50] This view led to a number of features not commonly observed in other 1970s-era operating systems, nor even in those existing today.

To ensure that the system met the needs of the reservations agent, it was designed to do a few specific things as quickly as possible. The operating system itself was small and lightweight, consuming little memory and CPU cycles for its own work. Its network control routines were highly-tuned, imposing less than five percent overhead on the CPU, compared to over thirty percent in other operating systems.[51] It also processed transactions in parallel using a cooperative "multiprogramming" approach.[52] Whenever an existing transaction needed data from a file or the network, ACP suspended that transaction and began processing new ones until the requested data was fully read and available in memory. This allowed ACP to process a large number of transactions simultaneously.

ACP was also designed to help the operators keep the system up and running. All normal maintenance could be performed while the system was still online. New versions of the application programs could be loaded dynamically without shutting down. All key system metrics could be constantly monitored, and all significant activity was both logged and streamed to a printer or terminal.[53] Most ACP installations used multiple, redundant CPUs, and ACP could

[48] Siwiec, p. 173.

[49] Knight, pp. 1424–1425.

[50] Siwiec, pp. 171–172.

[51] Knight, p. 1426. One IBM developer noted that this came with a certain risk—the system was not terribly protected, and one could easily write code that could bypass safeguards in order to obtain the necessary performance.

[52] Ibid., p. 1427. This kind of non-preemptive multitasking is usually quite dangerous, as programs only yield when they need data from a peripheral device. However, since PARS and BASE I transactions are short duration and always need to access peripherals, this approach was satisfactory for the time.

[53] This was also necessary in order to debug the system, as some logic errors may only occur under certain timing-dependent conditions. The logged information could also be analyzed to determine the actual usage of the system and tune it accordingly.

automatically switch to the secondary when the primary failed or needed maintenance. The database could also be mirrored onto redundant disks, allowing quick recovery. And since no system is ever perfect, ACP was designed to restart in a few seconds, as opposed to the multiple minutes required by other operating systems.[54]

These technical features were no doubt appealing, but perhaps the main selling point of ACP was that PARS provided a clear example of a working ACP system that was already operating at a scale far beyond BASE I. The original capacity requirement for BASE I was 3,500 authorizations per hour, and by 1976 this had climbed to 50,000, or nearly 14 per second; in contrast, the existing PARS installations were processing 50 to 75 transactions per second.[55] Furthermore, IBM's testing had shown that the transaction throughput was more or less linear with the CPU speed, suggesting that capacity would continue to grow as IBM developed faster processors.[56] BASE I was also much simpler than PARS, and required far less disk access per transaction.

The actual re-implementation of BASE I was accomplished by an IBM programming team familiar with ACP, and was completed in less than a year.[57] Fojtik stressed that this was almost an entire re-design of the system, as there was little documentation on the existing BASE I, and the ACP environment required a completely different approach. Once the software was ready, Peirce installed two IBM System/370 model 138 mainframes in the San Mateo center, one being a redundant backup for the other. The authorization messages were cutover to the new IBM system in late 1977, resulting in an estimated three-fold increase in capacity.[58]

Beyond the initial increases, moving to the IBM platform also allowed Visa to continue expanding their capacity without adjusting the software. The instruction set for the IBM System/370 mainframes always remained backwards compatible, allowing one to run software developed for an earlier model on a later, more powerful model without modification.

7.2.2 Multiple Data Centers

Although moving to IBM and ACP would greatly enhance the capacity and dependability of the BASE I installation at San Mateo, there was still only one data center, and all interchange authorization traffic flowed through it. While BASE I was being re-implemented on ACP, John Totten, who transferred to BASE I after BASE III was cancelled, expressed concern to Hock and Russell about this single point of failure:

> I remember going in for some work on the budgeting numbers in December 1976,

[54] Siwiec, pp. 171–172.

[55] BASE I numbers come from Tootelian and Peirce interviews. PARS numbers come from Siwiec, p. 172. For comparison, Visa claims that their average peak in February 2007 was 6,800 per second, and their peak capacity was 13,000 per second. See `http://corporate.visa.com/md/fs/corporate/visanet.jsp` (Accessed on 12 February 2007).

[56] Siwiec, p. 174.

[57] Peirce interview.

[58] Peirce and Schramm interviews.

> and I said, "... here we are, providing the service off of one computer in one data center, which is made out of wood, combustible wood, on a hillside that has dry grass, above a road where a car could catch on fire; we're right below a parking lot where kids are parking their cars, and the cars could come off the edge and drop into the building; and not only that, we're a mile from the San Andreas fault! How many more threats could you take on? If your goal is to run the credit card industry of the world, we really should have some sort of redundant, parallel site."
>
> I went back in [on Monday] and Dee said, "you know, we thought about what you said, and you're right. You have a new job. Your new job is to go somewhere on the East Coast, find a site and build a center." Nothing more. No papers. Nothing more than, "we thought about it, now let's do it." ... and Dee said, "Oh, and by the way, this has to be up by July 1977." So we had 6 months to find a site, build it, and staff it.[59]

Selecting the Site

Totten returned to the San Mateo center and set about establishing the criteria by which he would choose the site. First and foremost, they needed a highly-reliable and secure telecommunications service, so the new site had to be wired for the new digital communications system recently offered from AT&T. The new site also needed to be in a location where a suitable workforce would want to live. Totten toured the East Coast and came back with a recommendation. However, he neglected to consider the criteria that would be most important to Hock:

> The city we recommended was Charlotte, North Carolina, not McLean. Because McLean was much more expensive... so we came back and made the recommendation to Dee... and he said "Charlotte—who knows about that kind of place? I'm in the process of making a major move to take electronic banking into Europe, and I want to have the letter head, when someone in Europe sees it, to know that it's an important place. Now McLean, Virginia, is that near DC?" And I said, "yeah, it's right across the river." He said, "could we use a Washington DC address on our letterhead?" and Dave Hall [who joined Totten's team] said, "yes, I've already checked it out and there's a box, we've already signed up for, out at Dulles airport, which has a Washington DC address." And Dee said, "well that's it!"[60]

Hock was always conscious of Visa's image, not only with the public, but also within the banking industry. At this time, Visa's headquarters were in San Francisco, the banking capital of the West Coast, and Hock wanted Visa's presence on the East Coast to be in a similarly important place (presumably New York was not an option). Although Visa eventually moved its headquarters to San Mateo and then Foster City (smaller towns south of San Francisco), they still maintained a San Francisco mailing address by renting a post office box at the nearby San Francisco Airport.[61]

[59] Totten interview. Information on the creation of the East Coast data center comes from interviews with Totten, Peirce, Derman, and Fojtik.

[60] Totten interview.

[61] Cleveland, p. 47.

Totten selected a site in McLean that was on the same telephone grid as the CIA, something he felt would provide the security and reliability he wanted. He chose a steel-framed building that was still under construction, and leased the top two floors, which were then finished before the lower floors. When they moved in, the elevators were not even working yet, so the employees had to climb the exposed stairs. Totten remarked, "That was the typical way of Visa: we figured out every way to skin the cat to get it done quicker."

The new center was built and staffed by the July 1977 deadline, but Totten and his staff were still awaiting their redundant pair of IBM mainframes as well as the new BASE I software. Since they had a robust telecommunication system, they established a merchant authorization call center, similar to the one in San Mateo, providing backup and off-hours coverage for the acquirers in the eastern half of the country. Soon afterwards, the computers and the BASE I software arrived, and they set about installing the first dual-site ACP system.[62]

Designing a Dual-Switch ACP System

The main goal of creating the second data center was to provide a full, redundant backup system in another location. However, it would be somewhat wasteful to leave that system dormant until the primary system failed, which was the common practice for redundant mainframes within a single center. Instead, while the IBM programming team was re-implementing BASE I on ACP, they also added the features necessary to run multiple, cooperative systems in parallel. This was rather innovative at the time, as none of the airlines had ever run more than one concurrent ACP installation.

To accomplish this, a new telecommunication circuit was run from the McLean center to each end-point on the existing BASE I network. Each end-point had an affinity to a primary data center where it sent all of its outgoing authorization requests, but it could receive incoming authorizations from either center. If the primary center stopped responding, the end-point would automatically begin sending requests to the secondary center.

However, running parallel authorization systems required more than just switching the network. Recall that BASE I has the ability to stand-in for issuers and authorize transactions on their behalf when their systems are unavailable. The policies that dictate how much BASE I may approve over how long a period, the authorizations it approves under these conditions, as well as the card numbers that should be automatically denied, are all kept in the ACP database. These data, and changes to them, also needed to be replicated between the centers to ensure that no information was lost during an outage. To accommodate this, BASE I was enhanced to send data change notification messages between the various installations.

The dual-switch approach provided Visa with not only some added protection against regional threats, but also a convenient way to conduct major maintenance on a center without shutting off the entire service. If they needed to shut down the San Mateo center for any rea-

[62] Totten and Fojtik interviews.

son, all traffic could be switched to the McLean center in a matter of minutes. Either center was capable of handling the entire load on its own.

The dual-switch concept proved to be so successful that Visa eventually added major centers in England and Japan as well. All four centers run simultaneously, handling the authorizations for their given area, but the centers in the US and England can handle the entire world's traffic alone if needed.[63]

7.2.3 Expanding Internationally

International Authorizations

Moving BASE I to IBM/ACP and opening the second data center gave Visa the capacity it needed to handle the growing volume of domestic interchange authorizations. However, *international* authorizations were still quite slow and cumbersome. If a US cardholder made purchases in a foreign country, the acquirer needed to telex the issuing bank to obtain an authorization. The process was just as inefficient as domestic authorizations prior to BASE I, but suffered further due to the greater difference in time zones.

The ultimate goal was to expand the online computer network internationally, but in the 1970s this was not entirely practical, so Fojtik developed a rather ingenious stopgap. Building on a similar system he had done for Singer, Fojtik wrote some software that emulated a telex and provided a bridge to BASE I. When a foreign bank needed an authorization on a US card, they telexed a new number in San Mateo, which corresponded to a modem connected to one of the old PDP-11s. Fojtik's software then read the request, parsed it, reformatted it into a BASE I authorization request message, and submitted it to the switch. A few seconds later, his program received the response, which it then reformatted into a telex reply message. Because the operating regulations stipulated the proper ordering of the telexed information, it was rather easy to write the parsing software, and Fojtik remarked that it was actually quite forgiving. Because card numbers, expiration dates, and amounts were all distinctly recognizable, the software allowed them to be in almost any order with any amount of whitespace in between. Using this system, foreign acquirers could now obtain international authorizations within a few seconds, at any time.[64]

This "auto-telex" system was eventually replaced when the online computer network was expanded into other countries. The first links were to the UK and Canada in 1977, and within a decade, Visa had amassed enough leased lines and satellite links to connect every member bank and processor on the planet.[65]

[63] Sources indicated that a fifth center has been added in the US, but it is unclear if this new center will replace one of the existing ones.

[64] Fojtik interview.

[65] 'Visa Verification Net Reaches UK, Canada', *American Banker* (22 July 1977), p. 1.

Multi-Currency Clearing and Settlement

BASE II also expanded outside the US, though it was not until the mid 1980s that transactions were settled in multiple currencies, a necessary feature for the payment system to be considered truly "international."[66] From the early days of the BankAmericard licensing program, it was agreed that foreign transactions would be cleared and settled in US dollars only. The currency conversion was done by the acquirer, and although the operating regulations established certain limits on how and when this should be accomplished, most acquirers used this system to their own advantage.

In the late 1970s, the international members began using BASE II for electronic clearing and settlement, but transactions were still converted to US dollars by the acquirer, who chose the most opportune time to perform the conversion. They also typically increased the rate by a few percentage points in their favor to cover their conversion costs. For highly-volatile currencies, this could result in substantial profits for the acquirer. Cardholders were also confused when they received their bills, as the amount was not expressed in the local currency, and the dollar amount was converted at a different rate than the one in effect when the purchase was made.

The Eurocard system, which was affiliated with MasterCard in the US, began offering multi-currency settlement to their European members in the early 1980s, and MasterCard had announced that their US dollar members would also participate in that scheme starting in 1986. A team at Visa, headed by a foreign exchange expert named David Nordemann, was charged with developing a similar feature.

Nordemann devised an approach that would allow all members to clear and settle with the Visa system in one of many supported currencies.[67] Visa, as the central clearinghouse, would then perform all currency conversions using the wholesale exchange rates, plus a percentage fee that could be divided between the acquiring and issuing regions. This fee was primarily designed to compensate the acquirers who would now lose the income they gained from controlling the conversion rate. Issuers were also allowed to add a few percentage points to foreign transactions, provided it was allowed under local law.[68]

The Board accepted the proposal, and multi-currency settlement began in 1986. Implementing this in the BASE II software posed certain challenges, but finding a bank that would commit to fixed currency exchange rates each day was even more difficult. In order for multi-currency clearing and settlement to work, BASE II needed a set of conversion rates that would remain constant from the moment the data collection began to the final transfer of funds several hours later. During that time, the currency exchange rates continued to fluctuate, so any

[66] Information on multi-currency clearing and settlement comes from interviews with Nordemann and Schonheyder.

[67] Visa currently clears in 172 currencies and settles in 16 (`http://corporate.visa.com/md/fs/corporate/visanet.jsp`, accessed on 12 February 2007).

[68] Many issuers today charge between one and three percent for foreign currency transactions. Recent laws in the US now require that this be made explicit, as it was previously buried in the exchange rate reported on the bill.

bank willing to act as Visa's currency trader would be taking a significant risk. Eventually Barclay's, which had a large foreign exchange department, agreed to play this role, and was able to mitigate the risk by closely monitoring the markets and quoting rates based on their projections.

Multi-currency clearing and settlement is one of those quiet features that does not get the appreciation it truly deserves. Extending the authorization network internationally was certainly important, but allowing members to clear and settle in their own native currency established Visa as *the* premier worldwide payment system. With the addition of this feature, Hock's original vision of a global system for the exchange of value was nearly fulfilled.

7.3 Conclusion

In this chapter, we examined the various ways in which the payment system was expanded, both organizationally and technically, throughout the latter 1970s. The organizational expansions touched off an explosive period of growth for Visa, establishing them as the dominant bankcard system. However, to understand this period correctly, one must consider the changes to the organization and the computer systems together. The enhanced computer systems not only enabled Visa to handle the dramatic growth, they also made the entire payment system even more attractive to prospective members.

However, there was one aspect of the payment system that still fell short of Hock's expectations: the point of sale. Most merchants were still grappling with cumbersome hot card lists, paper sales drafts, and card imprinters. In order to make the Visa system a common, and ultimately preferred method of payment, they needed to eliminate the paper entirely, and capture the transaction electronically at the point of sale, even at the smallest of merchants. It is to this story that we now turn.

Chapter 8

Automating the Point of Sale: Encoding Standards and Merchant Dial Terminals

In the previous chapter, we saw how NBI/Visa expanded during the 1970s to become the dominant worldwide bankcard system. This was accomplished through a number of changes, both to the organization and the computer systems, which ignited an explosive period of growth for Visa, both in terms of sales volume and capacity.

In this chapter, we will examine the last area of the payment system to become fully automated: the point of sale (POS). Although a few banks began deploying POS terminals in the early 1970s, most merchants were still referencing dense lists of invalid card numbers, telephoning for verbal authorization, and manually completing paper sales drafts. Floor limits reduced the number of transactions requiring authorization, but they also allowed too much undetectable fraud and credit losses: an estimated one billion USD each year.[1] If the Visa payment system was ever to become a true replacement for cash and cheques, this first link in the payment chain needed to be automated fully, even at the smallest of merchants.

Hock's philosophy typically dictated that what happened between the merchant and the acquirer was their business, and Visa's influence on that relationship should be limited only to those aspects necessary to ensure the health and stability of the overall system. For most of the 1970s, Visa stayed out of the point of sale, concentrating instead on automating the exchange of information between members. This allowed the innovative banks to experiment with various kinds of POS terminals and card-encoding methods. As the banks committed significant amounts of time and money to their various solutions, intense debates broke out regarding which should become the national standard.

At the turn of the 1980s, Visa decided it was time to intervene, and forced a certain amount

[1] 'Visa Dial Terminal Pilot Project Final Report' (April 1982), p. 1.

of closure to these debates through two important actions. First, they mandated that all Visa cards issued after 1980 include a magnetic stripe encoded to their standard, effectively ending the card-encoding debate. Second, they encouraged the widespread use of POS terminals by stimulating equipment manufactures to build inexpensive devices that used standard voice telephone lines, and providing the economic incentives necessary for merchants to adopt them. These actions created the conditions by which full POS automation could be achieved at the national, and eventually international, level.

However, before we examine these two moves in detail, we must first briefly discuss the technical vision that informed them. This vision, and the various steps needed to achieve it, were developed during a project known as BASE IV.

8.1 Dreaming the Future: BASE IV

In 1974, after BASE II was put into production and the first version of BASE III was nearing completion, Hock began another systems development project that was his most ambitious yet. It was duly named BASE IV, but it also went by another name that was perhaps more descriptive for Hock: Electronic Value Exchange (EVE).

8.1.1 Electronic Value Exchange

Throughout the mid-1970s, the American banking industry was preoccupied with the possibilities of electronic funds transfer (EFT). Bankers were eager to replace cash and cheques with seemingly less expensive cards, terminals, and automated clearing houses (ACHs). ATMs also promised to reduce the need for tellers and their corresponding labor costs. From Hock's perspective however, the bankers were simply trying to automate the existing forms of banking rather than re-imagining how banking's central function, value exchange, should occur in an entirely electronic world. In his speeches, he exhorted bankers to think not just of EFT, but also the broader concept of EVE, lest they "may swiftly be hooting in the commercial graveyard where the ghosts of form, which did not follow function, are buried."[2]

BASE IV was Hock's own attempt at imagining a world in which every financial transaction was completed entirely in electronic form. In this world, transactions could originate from any device, be it a POS terminal, a cash dispensing machine, a pre-authorized transfer, or even a telephone banking interface.[3] Those transactions might access any asset the payor owns: not just a credit line, but also deposits, or even liquid investments. Similarly, those transactions might credit any asset the payee owns. In the middle would be a switching system that has

[2] Hock, 'Electronic Funds Transfer or Electronic Value Exchange?', p. 8.

[3] In the 1970s, the word "terminal" was often used for any kind of origination device. However, in order to avoid confusion, I will use the more distinctive terms employed today. By "POS terminal," I mean a device used to authorize (and possibly capture) transactions at the point of sale, and by "ATM" or "cash dispenser," I mean a device used to obtain currency.

connections to all the devices and all the assets, or at least to other networks that can reach those assets. This switching system, which would of course be built and operated by NBI/Visa, would connect everyone's assets together, making them accessible from any device, at any time. This, for Hock, was the essence of electronic value exchange.

IBM and Compata were hired to write the functional specifications for a system that could bring Hock's vision of EVE to life, and by 1975, their work filled a number of large binders. However, that was as concrete as the system ever became. When BASE III was cancelled, BASE IV was as well, for three principal reasons. First, most of BASE IV's components were intended to run in the member processing centers, and BASE III had demonstrated that NBI should not be in that business. Second, the necessary technologies were either unavailable at that time, or were far too expensive for the system to be economically feasible. Third, it was unclear whether the public was actually ready to abandon cash and cheques for electronic transactions. Peirce counseled Hock and Russell to "stick [the design] on the shelf and go on with something else, because you'll never be able to build this system. The world's not ready for this."[4]

8.1.2 Gems in the Rubble

Although BASE IV was never built, it still had a profound influence on the evolution of the core payment system. Peirce explained:

> The value of BASE IV was that it defined an end point and identified the building blocks and standards required to reach that end point. Almost everything we did subsequent to BASE IV was consistent with it.[5]

Two of Visa's subsequent actions—standardizing how the cards should be encoded, and stimulating the development and adoption of affordable POS terminals—will be discussed in detail in the following sections. However, there was another key idea generated during the BASE IV design that is worth mentioning here: a new kind of extensible message format. The original BASE I message format was fixed both in length and content, containing only the few fields necessary for credit authorization. This would simply not suffice in the all-electronic world envisioned for BASE IV. Multiple transaction types were already needed, and the payment and banking industries were still experiencing profound change, making it likely that entirely new, unforeseen products and services might arise in the future, requiring completely new types of transactions and message fields. As with any large, decentralized system, changes to the messages would need to occur gradually, and the format had to allow for some messages to include new fields while others omitted them.

[4] Peirce interview. Peirce stressed that as an IBM salesperson, he had every incentive to encourage NBI to pursue the plan, but he knew that it would end in disaster, and thus was unwilling to risk the long-term relationship between the two organizations.

[5] Peirce correspondence.

To accommodate these requirements, IBM developed a dynamic format that could be extended gradually over time without significant changes to the switching software. Each message began with a type indicator that distinguished between authorization requests, financial transfers, reversals, administrative messages, or any other kind of transaction that might be needed in the future. Following the type were eight bytes treated as a bitmap: each of the sixty-four bits corresponded to a data field defined in an external dictionary, and the value of the bit indicated if that field was present or absent in the message. To enable more than sixty-four fields, one of the bits was eventually reserved to indicate whether another bitmap and set of fields followed the current one.[6] Thus, messages could contain only the necessary and relevant fields, densely packed, with a relatively small eight-byte overhead per fieldset to indicate which fields were present.

As new fields were defined, the only change necessary to the switching software was a new entry in a field definition table. Although the standard allowed for variable-length fields, the values of such fields were always prefixed by their actual length, enabling the switching and logging systems to handle them opaquely. Issuers could also take advantage of new field values when present, or safely ignore them.

After a battle with the airline industry that is unfortunately too detailed to cover here, this format was eventually ratified as the standard for all electronic financial messages by the ANSI and ISO banking industry committees. The format, known as ISO 8583, allowed Visa to not only support new types of transactions over time, such as single-message debit, but also add new fields incrementally to existing transactions, such as the Card Verification Value (CVV) or the related three digit number in the signature panel used when the card is not swiped through a terminal (CVV2).[7]

8.2 Encoding the Card: Magnetic Stripes and Magic Middles

One of BASE IV's key design goals was to remove the paper sales drafts entirely from the system. All transactions were to be originated in electronic form at the point of sale, and that implied not only some sort of counter-top electronic device, but also a *machine-readable card*. A POS terminal that required manual entry of the card details would suffer from the same data entry errors that occurred before the use of embossed cards and imprinters. Thus, the POS terminals needed to "read" the card directly, and that implied a mechanism for encoding the account details onto the card itself in some sort of machine-readable form.

However, the BASE IV design did not specify any particular method for encoding the card, primarily because there was considerable debate at the time within the banking and airline industries as to how this should be done. Visa would eventually force a temporary closure of

[6] According to Derman, this chaining of bitmaps and fields was added during the ANSI standards process and was not part of the original IBM design.

[7] For a description of single-message debit transactions, see Chapter 9.

this debate, but to understand the issue fully, we must first review the various encoding options and the ways in which certain groups made claims about the superiority of their technique over others.

8.2.1 Encoding Options and Standards

In the early 1970s, there were two general approaches to making the card machine-readable: optics and magnetics. Within these two approaches, issuers experimented with various techniques, but they eventually standardized on Optical Character Recognition (OCR) and the magnetic stripe (often abbreviated as "magstripe"). The approach used generally depended upon the issuer's industry; the oil and retail industries normally preferred OCR, while the banking and airline industries typically favored the magstripe. However, this was not entirely universal, and as we shall see, one powerful bank developed and promoted a unique optical technique.

Simply defined, an OCR system "reads" alpha-numeric characters printed in specific fonts using optical sensors and shape-detecting algorithms. The retail and oil industries both made early commitments to this technology. Standard Oil of California installed what seems to be the first OCR-based data capture system for card sales drafts in 1956.[8] The National Retail Merchants Association formed an optical scanning standards committee in the late 1950s, which subsequently recommended the use of a similar system that could read product identifiers printed on labels.[9] In the 1960s, Addressograph-Multigraph perfected their "barcode" technique, which encoded the characters into a series of parallel vertical bars, and these proved easier for the machines to read reliably.[10] Most of the retail industry eventually adopted the barcode technique, especially the supermarkets, and electronic cash register manufactures began to build-in scanning wands to read them.[11]

Because the oil and retail industries were already using OCR to capture their payment card drafts, it was a relatively easy step to propose reading the card directly at the point of sale using a similar technique. The embossed characters on the card were already printed in an OCR-readable font so that they could be read from the imprinted draft, and the OCR sensors were beginning to reduce in size and cost.

The Data Source Corporation seems to be the first vendor to have developed an OCR-based, card-reading POS terminal. NDC, the large processor discussed in Chapter 5, installed these terminals at numerous service stations and retail locations in 1971, and the first unit installed reportedly caught an unauthorized card on its fourteenth transaction.[12] Data Source

[8] Herbert F Schantz, *The History of OCR: Optical Character Recognition* (Manchester Center, VT: Recognition Technologies Users Association, 1982), p. 11.

[9] Ibid., p. 13.

[10] Ibid., pp. 15–16.

[11] Campbell-Kelly and Aspray, pp. 176–180. In the early 1980s, Visa considered requiring a barcode on their debit card, which they were trying to sell to the supermarkets, but decided against it because it lacked the necessary capacity.

[12] 'NDC Credit Authorization Pilot Underway', *Payment Systems Newsletter* (July 1971), p. 7; 'Operations and Systems Notes', *American Banker* (16 February 1972), p. 6.

and NDC claimed that OCR was the best technique for POS terminals, as issuers did not need to add anything to their cards. However, they often neglected to mention that although the POS terminals could read the account number, which was embossed in the standard IMR-7B OCR font, they could not yet read the expiration date, which was inexplicably printed in a different font.

The banking and airline industries also made early commitments, but to magnetics instead of OCR. As noted in Chapter 3, the banks standardized on the Magnetic Ink Character Recognition (MICR) technique in 1957 for automating the processing of cheques and other bank drafts. MICR was never considered as a candidate for card encoding, but in the late 1960s, the airlines and IBM developed another magnetic technique that seemed promising: the magnetic stripe.[13]

The magstripe is essentially a piece of magnetic tape, similar to that used for computer data storage or audio recording, affixed by heat to a paper ticket or plastic card. Just like a computer storage tape, the magstripe can be encoded with binary data, which can then be extracted by passing the tape over a relatively inexpensive reader head. The airlines used this technique for their automated ticket vending systems, and adopted it as a standard for their own, privately-issued payment cards in 1969. The various cash dispenser manufacturers also adopted this technique for their access cards in the early 1970s.

As discussed in Chapter 5, many banks conducted POS terminal tests using the magstripe during 1971. When plans for these tests became known, the American Bankers Association (ABA) formed a task force to develop standards for card encoding, hoping to avoid unnecessary duplication and fragmentation in the marketplace.

The ABA task force evaluated a few different encoding techniques according to four basic criteria: availability, reliability, cost, and security.[14] In MacKenzie's terms, these criteria were four "properties" of the artifact, about which the task force would construct knowledge and disseminate it via the authority of the ABA.[15]

In their report, they concluded that the magstripe "offered the greatest opportunity to satisfy a wide variety of requirements for both present and future needs."[16] In many ways, their conclusion was entirely expected, as most of the task force members came from those banks already planning POS terminal tests using magstripes. Later critics would question whether the task force members recommended the magstripe primarily because the members had already

[13] It may be surprising that MICR was never considered, but Perry Hudson, who served as chairman of the ABA card standardization task force, remarked that it was simply never suggested (Hudson interview).

[14] Hudson noted that OCR was the only serious competitor to the stripe, but at the time of their investigation (1970), OCR POS terminals were not yet available. Additionally, Magtek was already selling a magstripe POS reader that was being used by the airlines.

[15] MacKenzie argues that we construct knowledge about the properties of artifacts according to four methods: testimony of authority; induction through testing; induction through use; and deduction. Each of these methods involves a social dynamic, even deduction, which is typically thought to be asocial. See Donald MacKenzie, 'How Do We Know the Properties of Artefacts? Applying the Sociology of Knowledge to Technology', in: Robert Fox, editor, *Technological Change: Methods and Themes in the History of Technology* (Amsterdam: Harwood, 1996).

[16] 'Magnetic Stripe For Credit Cards Urged by ABA Unit', *American Banker* (16 February 1971), p. 1.

made large investments in it.[17] However, this accusation had an obvious defense: those who had already made investments in the magstripe would have done so only after their own investigations, and their conclusions were unlikely to change when they made the same investigation as a member of the task force just a few months later.

The report offered several reasons why the magstripe was recommended over other techniques. It was already widely available, and as an established technology, had proven its reliability and durability in the field. Although adding the stripe virtually doubled the cost of the card, the reader heads were less expensive, simpler, and more reliable than OCR sensors. The stripe also offered a large capacity, allowing the inclusion of data beyond the basic account information visible on the card. Finally, the stripe was volatile, allowing for the possibility of devices that could write back to the stripe, which might be desirable for off-line systems.

In terms of availability, reliability, and cost, the magstripe genuinely seemed to be the best option, and in the subsequent debates, these points remained largely uncontested. However, the final criterion of security was not so straightforward. Regarding this property, the task force originally wrote:

> While no encoding technology was thought to be foolproof against fraud, the magnetic stripe was thought to have the greatest security against casual fraud because it is difficult to alter, the data is not visible, and it requires a fairly high level of sophistication and collusion to counterfeit.[18]

As we shall see in section 8.2.3, all of these claims, as well as the understanding of *where* security should be evaluated, would soon be contested.

8.2.2 Magstripe Tracks

Although the airlines and the ABA recommended the magstripe for their cards, they did not agree as to how the card information should be encoded upon it. Of central concern was the density at which the data should be recorded. Magnetic tape contains a large number of contiguous ferrite-oxide particles, and it is somewhat arbitrary as to how one divides them into discrete segments representing binary values. The technology in general use at the time was able to read reliably at a density of 75 bits per inch (bpi), but newer equipment in testing promised to read just as reliably at 210 bpi.

The ABA wanted to establish a standard quickly because the various pilot tests were scheduled to begin in the near future. Because the POS terminals being used in these tests would perform online authorization only, the banks needed only two fields on the magstripe: a thirteen-digit account number and a four-digit expiration date.[19] Using a 5-bit per character encoding

[17] For example, see Phillip Brooke, 'ABA Magnetic Stripe Highly Vulnerable to Fraud', *American Banker* (3 November 1971), p. 6.

[18] Quoted in 'Magnetic Stripe For Credit Cards Urged'.

[19] At this time, the account number was actually variable in length, but tended to be around thirteen digits. Sixteen digit numbers were not used until later.

at 75 bpi would provide the banks with space for 40 numeric characters, which was more than enough for their needs.[20] Additionally, they felt that the lower density would be more durable, as the lighter packing created a lower potential for bit-dropout, a condition often caused from damage incurred by the many unorthodox activities for which the cards were often used, such as scraping ice off windshields or opening locked doors.[21]

However, the airlines required more than the account number and expiration date. To enable fully-automated ticketing, the magstripe also needed to carry the cardholder's name. Names require alphanumeric characters, which implied at least a 7-bit per character encoding. A density of 75 bpi simply did not provide enough characters, so the airlines were willing to wait until 210 bpi readers were commercially available.[22]

The issue was eventually resolved by subdividing the stripe into two tracks, the first recorded at 210 bpi and the second at 75. The airlines claimed the first track and defined a field layout that included the cardholder name, account number, expiration, and a few airline-specific fields. The ABA took the second track and defined a field layout that included only the account number and expiration. The magstripe on all cards would contain both tracks, and devices could read either or both. To this day, payment card magstripes still contain two tracks, encoded at different densities, containing roughly the same information.[23]

Soon after this compromise, a third track at 210 bpi was also defined. Recall that one of the benefits of the magstripe cited by the ABA task force was its changeability. The ABA allocated this track primarily for read/write scenarios, where terminals and cash dispensing units would alter the contents based on the last transaction. However, the definition of the track was given over to the thrift industry, and although they recognized the value of writability, they saw it more as a chance to encode sophisticated personal identification data, such as a digitized finger or voice print.[24] The third track was briefly used in a read/write mode for off-line cash dispensing units, but was quickly abandoned after bankers realized just how easily it could be read and altered using widely-available audio equipment. As telecommunication costs in the US dropped, and ATMs and other terminal devices became strictly online, most issuers removed the third track in order to reduce the height of the stripe.

[20] The encoding scheme used 4 data bits plus one parity bit per character, allowing for 16 distinct values. This was enough for the digits 0 through 9, plus a few special values used as field separators and begin/end markers.

[21] This point was made by Bertram Tobin of Chase Manhattan Bank, quoted in 'Standards are the Glue', *Payment Systems Newsletter* (July 1972), p. 1.

[22] The airline encoding scheme used 6 data bits plus one parity bit per character, providing a maximum of 79 alphanumeric characters.

[23] Derman interview. Note that the format of these tracks was slightly redefined by Visa. See section 8.2.5 starting on page 155.

[24] Phillip Brooke, 'Completion of Standards for Thrifts Seen Spurring Conflict Over Magnetic Stripes', *American Banker* (14 September 1973), p. 1.

8.2.3 Magstripe Security

Shortly after the ABA task force defined the encoding standards for the second track, George Warfel from the Western States Bankcard Association began questioning its security, pitting his authority as a trained engineer against that of the ABA. He did not mince words: "Upon encodement with the American Bankers Association-proposed format, you create a fraudable document that can plague the entire credit card community. . . ."[25] He revealed that for about $150, an amateur can build a fairly simple device using widely-available audio parts and transistors that could copy the contents of one ABA-encoded stripe to another. He called it a "skimmer," and noted that "It is just like the tape-dubbing machine used by college students to copy tape cassettes—except it works for cards."

Playing on the fear of disreputable teenagers, he presented a scenario where a young gas station attendant could use one of these devices to copy the data from the stripe of a legitimate card to the stripe of a stolen one. Because POS terminals read only the stripe on the stolen card, and not the embossed numbers on the front, any typical transaction the teenager made with the stolen card would be authorized. However, sales drafts generated using a card imprinter would pick up the embossed numbers, which were for an account that was no longer valid.[26] At this pre-BASE II time, it might take weeks before that draft reached the issuer, and only then would the issuer realize that the stolen card now had a new stripe. By the time the issuer could determine which account had been skimmed, the attendant could easily skim a different card, creating a near endless cycle of unstoppable fraud. Furthermore, if the stripes and terminals were shared between the different payment networks, one could easily skim the stripe from a BankAmericard and copy it to a Master Charge or American Express; it would likely take months for the card organizations to determine what had happened.

To be clear, Warfel's concern was not with the magstripe itself, but with the way the ABA proposed to encode it, and the environment in which it would be used. Warfel observed that the only environment in which the magstripe was currently used was cash dispensing. The access cards for these devices employed magstripes, but they were encoded and used in much more secure ways. For example, Docutel used four tracks to hold a "deviously scrambled code" that must be input by the cardholder to complete the transaction. Burroughs included a second, unalterable stripe that was encoded at the time of manufacture. IBM used a forty-five degree rotation on their more sophisticated reader heads. The cash dispenser manufacturers could do all this because their devices were used under controlled conditions, and the added encoding security contributed little to the $15,000 to $25,000 cost of a dispenser.

In contrast, the ABA-encoded stripe would be included on millions of cards and used in hundreds of thousands of POS terminals. In order for this to be economically feasible, the

[25] Phillip Brooke, 'ABA Magnetic Stripe Highly Vulnerable to Fraud', *American Banker* (3 November 1971), p. 6. This article is an edited version of Warfel's speech to the Data Processing Supplies Association.

[26] Recall that these early POS terminals performed authorization only. Merchants still completed a paper sales draft for each transaction, and it was these paper drafts that would be cleared and settled.

ABA proposed a format that was much simpler and did not require any additional user input. But Warfel felt that the ABA went too far, suggesting that their format was the "extreme in simplicity" and was thus vulnerable to unimaginable amounts of fraud. Although the format may have been adequate a few years earlier, technology had developed in the meantime to the point where any hobbyist could build a skimmer:

> Today you can buy transistors like gum-drops, in plastic bags by the dozen. The radio catalogs list 'credit card reader heads' along with hi-fi components. Thus, what was secret in '67 is common knowledge today. What was secure in '67 is vulnerable today.[27]

Given the vulnerability of the ABA-proposed stripe, Warfel questioned whether the banking industry should adopt OCR instead. By this time, Data Source was offering their OCR terminal, and since it read the embossed numbers on the front, the authorization and sales draft would always refer to the same account. This was not completely secure though, as the numbers could be cut off and re-arranged, or melted down and re-embossed, but it would be easier to detect than a skimmed magstripe. However, Warfel realized that switching to OCR may not be possible given the existing investments in magstripes, so he also suggested that the ABA at least require the use of high-coercivity materials. Once encoded, this kind of stripe was essentially unalterable by a hobbyist, as changing it required special writing heads and high amounts of magnetic energy.

The ABA task force quickly responded to Warfel's criticisms by acknowledging the vulnerabilities of the magstripe, but arguing that with some minor adjustments, such as the use of high-coercivity materials, the magstripe was still more secure than any other available option.[28] They urged the banking industry to wait for the results of the various POS terminal tests being conducted by the banks, as these would reveal not only how secure the stripe was in practice, but also ways in which it could be made more secure.

This was a subtle but important shift in the task force's approach. They now asserted that the proper level at which security should be discussed is the *entire payment system*, not just the particular card encoding technique. In this way, they were redefining the locus of the security property, focusing the participants away from the magstripe artifact towards the system as a whole. The task force was envisioning a future environment where the POS terminals would not only authorize transactions, but also capture and transmit them electronically for clearing. Thus the stripe would become the sole location of account data, eliminating the central problem Warfel described. Even if a card was skimmed, the task force believed that fraud detection algorithms running at the issuer or on the switches would quickly detect it, and a simple examination of the transaction logs for the affected accounts would reveal the merchant location where the skimming occurred.

[27] Warfel quoted in Brooke, 'ABA Magnetic Stripe Highly Vulnerable to Fraud'.

[28] Phillip Brooke, 'ABA Defends Card Stripe Against Fraud Ease Charge', *American Banker* (3 November 1971), p. 1.

The debate subsided for nearly two years as various issuers conducted their POS terminal tests, but it resurfaced again in April 1973. Jack Scantlin, chairman of Transaction Technology Incorporated (TTI), a subsidiary of Citicorp, had been "suspicious of magnetics technology for quite some time," and decided to see just how easy it was to compromise the ABA-encoded stripe. Scantlin invited twenty-two students from the California Institute of Technology (Cal Tech) to design devices that posed a serious fraud potential, promising an attractive $5,000, $2,500, and $1,000 for the three best entries. Each team submitted their devices, along with a detailed report, and although TTI refused to discuss the devices for fear of giving criminals too many good ideas, *Business Week* described two types of skimmers built for as little as $25.[29] Scantlin then issued a press release announcing that he had discovered a "cheap and easy way to defraud the magnetic stripe," and called into question the ABA's commitment to the technique.[30]

The ABA's response to Scatlin was the same as their response to Warfel. However, this time, the ABA questioned the actions of TTI, as its sibling Citibank was already represented on the ABA task force. If TTI or Citibank had any exceptions to the magstripe, or any suggestions for improving it, they could have raised them at one of the task force meetings. Instead, they chose to issue a press release designed to characterize the magstripe as the "Achilles heel of the whole credit card system."[31] In response, Citicorp attempted to distance themselves from TTI's actions, stating that corporate management was unaware of the contest and would have never condoned it.

When asked why they encouraged students to develop devices capable of generating fraud, John Reed, the Executive Vice President of Operations for Citibank replied "we simply wanted to know about card reading technology." However, unlike Warfel, TTI and Citibank had an economic motivation for questioning the security of the magstripe. They had recently invested more than $30 million developing a different kind of proprietary, non-magnetic encoding technique. They were also just about to use it for nearly a million Citicards, and would soon offer it to other banks on a licensing basis. They called it the "Magic Middle."

8.2.4 The Magic Middle

Citibank and TTI never publicly revealed the technical details about the Magic Middle, but much can be deduced from the patents granted to TTI in 1972 and 1973.[32] The Magic Middle was essentially an optical form of a computer punch-card. As the name indicates, a payment card using this technique contained a special middle layer, sandwiched between the normal top and bottom layers of plastic. The outer layers were opaque to the human eye, but infrared light

[29] 'Beating the New Credit Cards', *Business Week* (11 August 1973), p. 120–122.

[30] Phillip Brooke, 'Citicorp Says Students Find Cheap, Easy Ways to Defraud Magnetic Stripe Cards', *American Banker* (9 April 1973), p. 1.

[31] Phillip Brooke, 'Bank Credit Card Leaders Back Magnetic Stripe', *American Banker* (13 April 1973), p. 1.

[32] For the basic concept, see United States Patent number 3819910. For details of the actual encoding method, see numbers 3775755 and 3858032.

could shine through them. The middle layer was made of a material that would reflect infrared light, except for the places where holes were punched through it. Thus, a reader could shine infrared light from one side, and a sensor on the other side could detect light coming through the holes, determining which were punched and which were not.

The middle layer contained four horizontal data tracks and one clocking track in the center.[33] The tracks were read vertically, one column per character, allowing for a 4-bit encoding scheme. The maximum capacity on a standard payment card was not disclosed, but since it was used by Citibank for many years, it was likely as much as the lower-density ABA track on the magstripe. The tracks ran across the width of the card, below the magnetic stripe and embossed account numbers, allowing it to coexist with the other encoding methods.

Armed with the results of their Cal Tech contest, Citibank and TTI claimed that the Magic Middle was far more secure than the magstripe for two reasons. First, the encoding technology was not even visible to the human eye, so it had a measure of security through obscurity. Second, even if the technique became well-known, it could not be copied without highly-specialized equipment.[34] Of course, this also implied that the card was far more expensive to manufacture, but Citibank argued that this was a small price to pay for a completely secure encoding technique.

Citibank and TTI christened the Magic Middle in October 1973, after they had already issued nearly a million Citicards containing it and installed terminals at 1,200 merchants that could read it.[35] Customers could use their Citicards for purchases at those merchant locations, in Citibank ATMs, and for identification and cheque cashing at any of Citibank's 226 branches. Two years later, Citibank would also install special interactive terminals, called "Citicard Centers," in their branches that would enable cardholders to complete basic account management tasks without waiting for, or speaking to, a teller. Interestingly, they admitted the Citicard Centers were "designed to begin conditioning customers to use the terminals and to allay their fears of computer technology."[36]

When the Magic Middle was announced in 1973, Citibank articulated a non-confrontational strategy regarding industry standards. Their spokesperson said:

> Realistically, it is not within our plans to convert the whole industry. We feel we have a better mousetrap than the mag stripe and embossed systems. We honestly feel this is a better way. Whether our technology dominates is not relevant to us. It was developed for our own use.[37]

[33] The clocking track had every other row punched, providing a simple means for ensuring proper alignment.

[34] To copy a Magic Middle card, one would need to obtain material that would reflect infrared light, extract the middle layer from a valid card, and use that as a template for punching the same holes into the blank material. However, to extract the middle layer, the thief would need to steal the legitimate card, which would likely be disabled as soon as it was reported missing. In contrast, skimming a magstripe could be done without the cardholder's knowledge, so the account would remain valid for much longer.

[35] David O Tyson, 'Citibank Introduces Check Card Different From Others', *American Banker* (25 October 1973), p. 1.

[36] Phillip Brooke, 'TTI: Citicorp's Own EFT Think Tank', *American Banker* (8 May 1975), p. 1.

[37] Mark Ponton, VP of Marketing, Personal Banking Group, quoted in Tyson, 'Citibank Introduces Check Card'.

However, Citibank changed its position less than a year later, announcing that they would make the Magic Middle available for licensing nationwide.[38] Furthermore, they began working with National Cash Register (NCR) and Docutel to add Magic Middle readers to their electronic registers and cash dispensing machines. Citibank was now directly challenging the *de jure* ABA magstripe standard, hoping to establish their own proprietary technology as the new *de facto* standard.[39]

The ABA responded to Citibank and the Magic Middle by simply adjusting the card encoding requirements to exclude proprietary technologies.[40] In a new amendment to their report, they required that any encoding technique proposed as a standard must not incur licensing fees; have sufficient capacity for recording all necessary information; and be available to any card issuer regardless of industry. The ABA spokesperson admitted that these requirements effectively disqualified both the Magic Middle and OCR, but assured that they were in the best interest of all issuers.

Citibank continued with its licensing plan, but was unable to sign enough banks to pose a serious threat. In many ways, this was not surprising, as few banks would be willing to pay the licensing fees and higher manufacturing costs, not to mention trusting what Spencer Nilson called "the world's most hated bank" with their encoding technology.[41] When the major West Coast banks, such as BofA and Wells Fargo, ventured into POS terminals, they chose the new devices from Data Source instead, which could read both the magstripe and the embossed account numbers and compare them to detect skimming.

8.2.5 Visa and the Magstripe

Throughout all of these debates, NBI/Visa remained mostly a passive observer, occasionally commenting that although the magstripe seemed likely to become the accepted standard, they would not rule out the emergence of a better encoding technique.[42] As the 1970s drew to a close, Hock decided that it was time to encourage full automation of the point of sale. This would not only bring Visa one step closer to the all-electronic value-exchange system he envisioned in BASE IV, it would also enable a zero floor limit environment, which would dramatically reduce the effects of fraud.[43] However, this also implied that Visa needed to mandate a standard encoding technique for all cards bearing their mark, both domestically and internationally.

[38] Phillip Brooke, 'National Licensing of Citicard Planned', *American Banker* (11 September 1974), p. 1.

[39] To clarify, the ABA actually has no legal or political power to force banks to follow their standard, but they are recognized as an authority within the banking industry.

[40] 'ABA Issues Credit Card Criteria, Reaffirms Support of Magnetic Stripe', *American Banker* (14 November 1973), p. 1.

[41] 'Citibank Continues Out-of-State Credit Card Promotions', *The Nilson Report* 182, p. 1.

[42] For example, see Phillip Brooke, 'Magnetic Stripe Credit Card Encoding Nears Worldwide Test', *American Banker* (18 October 1972), p. 1.

[43] In a zero floor limit environment, every transaction is authorized, which allows issuers to identify and stem fraud quickly.

The task of establishing Visa's encoding standard fell to Win Derman, who had already been working with the various standards groups on the bitmap message format.[44] In addition to the basic ABA-defined fields, Derman defined two new ones that would be needed in the Visa context. The first was a three-digit "service code," which at the time, merely indicated if the card could be used in international interchange, domestic interchange, or no interchange at all. Some members of Visa International operated in non-exchangeable currencies, so the magstripe needed to include information that would prohibit use outside the country of issue. Although one digit would have sufficed at the time, Derman argued that they should reserve two more for future needs, and today, these are used to indicate what kind of authorization is required, and what services the card supports. The second field was a PIN verification value (PVV), which is generated through a one-way encryption algorithm. The PVV allows a terminal or ATM to verify a PIN in an off-line environment.

Although these two fields were small-enough to fit into the remaining space on the ABA-defined second track, Derman also wanted to take advantage of the larger capacity of the first track. By this time, 210 bpi magnetic reader heads had become standard, and the ABA, wanting to take advantage of the larger capacity, had unfortunately established a banking-specific format for the first track that differed from that used by the airlines. To distinguish the two, both industries agreed to add a single-character format code to the front of the stripe: A for airline and B for banking.

Derman wanted to build upon the airline's format so that Visa cards could be read easily by airline POS terminals. During this period, Visa was also concentrating on improving their presence in the T&E market, so compatibility with the airlines' systems, which American Express already had, was paramount. However, convincing the member banks to switch to the airline format for the first track would be difficult, as many were already issuing cards using the banking format.

His solution was to recast the debate as a domestic versus international one, knowing that Visa was now an international organization, and any standard they proposed would need to be seen as international as well. Fortunately for Derman, when the ABA defined their format for the first track, they encoded the expiration date in the American MMYY format.[45] The airline format, which was established through the International Air Transport Association (IATA), used the more international form of YYMM. Since the ABA recommended standards only for the US, international banks also tended to use IATA's YYMM format.

Derman presented his case to the US banks, arguing that the ABA format for the first track was US-centric, and since Visa was an international organization, their encoding standard should use an international date layout. Since the airline format already expressed the expiration date in an international manner, it would make sense to build upon that. The existing format code would then allow the US banks to gradually transition to the airline format, as

[44] Information on the Visa magstripe standard comes from interviews with Derman.

[45] That is, a two-digit month followed by a two digit year.

the POS terminals could still read the banking format during the changeover.

The US banks eventually agreed, and Derman got the standard ratified by Interbank, the ABA, ANSI and ISO. The banks and the airlines were now finally on a common standard for the first track, although Visa cards would also continue to carry the now-extended second track. In 1979, Visa adopted a bylaw requiring all cards bearing their mark issued after 1980 to include a magnetic stripe, encoded according to the Visa standard.[46] Citibank, which had joined Visa after the ban on dual membership was lifted, threatened to sue Visa for effectively destroying their potential licensing market. However, they ultimately capitulated in exchange for a three month extension of the magstripe deadline.

Visa's mandate of the magstripe forced a temporary closure to the card encoding debate. It would resurface again when card manufacturers perfected the embedding of a computer chip in plastic, but for now, the path towards the development and mass-adoption of inexpensive POS terminals was paved.

8.3 Dialing for Dollars: The Merchant Dial Terminal Project

Now that all Visa cards would carry a standardized magstripe, the next step was to create the right incentives for the mass-adoption of POS terminals. As already discussed, a few different manufacturers were offering authorization terminals at this time, but these devices reportedly cost between $1,000 and $2,000 each, making them far too expensive for low-volume merchants.[47]

Although the unit costs were high, the real expense was in the way these early terminals communicated with the acquiring processor, as they required separate, dedicated leased lines, which incurred hefty monthly fees. For merchants who processed only a few card transactions a day, the fixed-price of the leased line combined with the merchant discount seriously reduced their profit margins. If Visa was to bring about the all-electronic BASE IV world, they needed to stimulate the development of a new kind of terminal that every merchant could afford to purchase and use, and then provide the right economic incentives to encourage merchants to adopt them.

8.3.1 Dial Terminals

In 1979, Visa started a new project to define the requirements and functional specifications for these affordable terminals. Technical leadership of the project fell to Frank Fojtik, who had previously spent time thinking about terminals when he joined NBI in 1974. Given his prior experience in telecommunications and electronic cash registers, Fojtik realized that the

[46] Jeffrey Kutler, 'Visa International Plans Magnetic Stripe For All Cards', *American Banker* (6 February 1979), p. 1. See also Jeffrey Kutler, 'Visa Approves Interchange Stripes', *American Banker* (12 June 1979), p. 3.

[47] Information on the Dial Terminal Project comes from interviews with Fojtik, Harrison, Derman, Powar, Peirce, and Pittenger, as well as the 'Visa Dial Terminal Pilot Project Final Report' produced by Visa USA (April 1982).

best option for low-volume merchants would be a terminal that could use their existing voice telephone lines. Any merchant accepting cards for transactions over the floor limit was already using that line to call for authorizations, and if the terminal dialed a similar number, it would incur no additional cost. In fact, the terminal should obtain the authorization more quickly, resulting in a shorter call time.

Visa briefly considered producing the dial terminals themselves, but ultimately realized that they lacked the necessary expertise. Furthermore, manufacturing POS terminals was not what they considered their core purpose to be, and their foray into writing card processing software (BASE III) had convinced them not to stray from that again. Instead, they decided to develop the devices in cooperation with multiple vendors. Visa would establish the requirements and write the detailed functional specifications, and the assorted vendors would design and build compliant terminals, competing with one another for what promised to be a sizable market.

The requirements and specifications were kept rather simple. The terminals would perform authorization only, so there was no requirement for internal data storage or an integrated printer. The terminal must have a small keypad with which the merchant can enter the purchase amount and send the authorization request. It must have some method for displaying the response and authorization code to the merchant. It needed to be as small as possible, as merchant counter space is precious, and it needed to run on a conventional power supply. Lastly, and most importantly, merchants must be able to obtain the resulting device for less than $500.

8.3.2 The Pilot Test

Visa found four manufacturers willing to build such terminals for a pilot test: GTE, Northern Telecom, Sweda International, and Taltek Electronics.[48] A total of 800 units were ordered from these vendors, and Visa USA members installed them at various types of merchants in disparate geographical regions. Besides making standard local calls to an acquiring processor, some terminals were configured to call a machine in Visa's San Mateo data center to test inbound Wide Area Telephone Service (WATS) lines, and others called local access nodes for the relatively new GTE Telenet packet switching network.[49]

IBM developed the software to answer calls placed to the San Mateo center and transform the terminal requests into BASE I authorization messages. This software ran on IBM Series/1 computers, which would soon be deployed at the member processing centers to replace the DEC Tape Transmission Units used by BASE II. These powerful minicomputers, known at the time as "Member Interface Processors" (MIPs), were also employed as gateways to BASE I, so it was relatively easy to interface the dial terminals to the authorization system through those

[48] 'Dial Terminal Pilot Project Final Report', p. 3. By the end of the pilot test, several more manufacturers had developed terminals based on Visa's specifications, including the now famous Verifone.

[49] Although the final report mentions local calls to acquirers as one of the telecommunication paths, Powar recalled that all calls during the test actually came to San Mateo.

machines.[50] After the pilot test, this software was made available as a standard package on the MIP, enabling the POS terminals to call the local acquiring processor rather than the San Mateo data center.

The pilot test ran from June 1980 to December 1981, and it proved to be a success on nearly every front. From a technical standpoint, the terminals worked exceedingly well, and merchants reported that they were almost never inoperative. Visa's requirements stipulated a minimum mean time to failure of eighteen months, but the average in practice was nearly thirty-six months. The telecommunication performance was more mixed: terminals that made local and intrastate WATS calls were reliable and cost effective, but interstate WATS was too expensive due to tariff increases, and the GTE Telenet service experienced operational problems. The final report noted the need for interstate WATS would be eliminated once the answering software was made available on the MIP, and GTE was steadily improving their reliability.[51] Despite using dial-up connections, the speed of the system was also adequate. An authorization made through the terminals typically took twenty seconds, which was at least twice as fast as one obtained verbally. It also involved less manual data entry, and thus allowed fewer opportunity for errors.

From a business and fraud-control perspective, the pilot test results were even more impressive. The creation of BASE I had helped limit fraudulent transactions, but even as late as 1980, only twelve to fifteen percent of interchange transactions were actually authorized—the rest fell under the merchant's floor limit. In contrast, merchants participating in the pilot test authorized every transaction, and within the first few days, the number of recovered cards rose sharply. Over the duration of the test, more than 3,000 cards were recovered by participating merchants, and over 10,000 transactions were declined that would have been allowed under the existing procedures.[52] The terminals also served as a powerful deterrent; one member bank reported that "...incidents have already occurred in which a thief, upon seeing his stolen card being put through the terminal, has turned and literally run from the store."[53]

The terminal's ability to deter fraud benefitted not only the issuers, but also the merchants. For transactions under the floor limit, merchants were supposed to check the card number against a large, dense list printed in a warning bulletin, but many found this arduous, time-consuming, and potentially insulting to the customer. If the number was in the warning bulletin and the transaction proved to be fraudulent, the issuer could levy a chargeback against the merchant, resulting in a deduction from the merchant's account for the transaction amount plus

[50] The MIPs were later renamed "Visa Access Points" (VAPs). Powar noted that they used the limited number of MIPs available that year as an incentive for terminal adoption; to get a MIP, the member had to purchase 200 terminals.

[51] Interestingly, Visa later used the CompuServe network as an inexpensive way to route authorizations from dial terminals to BASE I. Since CompuServe offered local access numbers in nearly every US city, the dial terminals rarely needed to make long-distance calls. This route quickly became the preferred one, primarily because CompuServe maintained the local telecommunications infrastructure (Pittenger and Powar interviews).

[52] 'Dial Terminal Pilot Project Final Report', p. 3.

[53] Ibid., p. 37.

a ten dollar handling fee. Merchants participating in the pilot test simply swiped every card through the terminal, resulting in an eighty-two percent decrease in their chargebacks.

8.3.3 Adoption Incentives

Although a promised reduction in chargebacks might have been attractive enough, Visa USA realized that they needed to offer merchants a stronger, positive incentive to adopt the terminals. Their solution was a reduction in the interchange reimbursement fee, the powerful economic setting used to balance the issuing and acquiring sides of the system.[54] Transactions authorized through POS terminals would qualify for a special Terminal Interchange Reimbursement Fee (TIRF) of one percent, which was twenty-percent less than the current average. Since this is the fee acquirers pay issuers, the TIRF provided an incentive for acquirers to push the terminals, which they did by reducing the merchant discount fee by a similar (though perhaps not equal) amount. For merchants, the lower discount fee, reduction in chargebacks, increase in authorization speed, and inexpensive unit price, all combined to make the terminals a net benefit.[55]

The TIRF provided the economic incentive necessary for mass-adoption of the terminals, but obtaining Board approval for such a reduction would be difficult. The interchange fee was an important source of revenue for issuers and the large banks were reluctant to fund automation that would primarily help acquirers and merchants. For these larger issuers, the warning bulletin was sufficient, as it provided a convenient way to justify chargebacks.

This difference in perspectives created a palpable tension. Hock saw terminals as an important step towards a totally electronic value exchange system, but the large issuers were focussing on their individual programs. Visa's CFO recalled that,

> Dee wanted to increase the Visa momentum and deliver the "knock out punch" to all of his competitors and detractors... Dee had the industry lead, but he wanted everyone else to believe that Visa had all of the answers, to choose anyone else would be suicidal.
>
> But others were growing tired of Dee's... big ideas.... As for the typical banker stereotype of the period, they were happy the system was "fixed," they were all making money, they all wished deregulation and competition would go away, and now they wanted to get back to playing golf, not embark on the next grand plan that Dee had devised.[56]

This seems to have been a common theme throughout Visa's early history: Hock often had to resort to manipulation or tricks to get Board members to put aside their own self-interests in

[54] The importance of this fee is discussed in section 4.4.

[55] The TIRF was not the first time Visa USA had offered a special incentive rate. In 1979, they offered an Electronic Interchange Reimbursement Fee (EIRF) to merchants who used their electronic cash register systems to authorize all transactions and submit them electronically to their processor within three days. This rate was originally developed to entice JC Penney, one of the three largest retail chains in America, to accept Visa cards. The details of this deal, which became highly controversial, will be discussed in the next chapter.

[56] Cleveland, p. 20.

order to improve the overall system. Although these improvements invariably benefitted those who initially opposed them, it was difficult to convince those members of such ahead of time.

8.3.4 Resolutions

Hock's plan was to introduce a resolution that would commit the Visa USA members to authorizing eighty percent of all interchange transactions within five years. Achieving such a goal would require the mass-adoption of terminals, which in turn would require the TIRF. Hock knew this would be a controversial proposal, so he purposely omitted it from the proposed agenda mailed to the Directors before the Board meeting.[57]

The meeting for that year, 1982, was held in Bermuda, and in keeping with previous years, Visa invited and paid for spouses to attend. This was advantageous for a number of reasons. First, and most important, the spouses were actually invited to attend the Board meetings, primarily because, as Visa's CFO put it, the Directors "would be more inclined to act civil, consider the good of Visa over their bank's self interest, and in other words, vote for whatever Dee wanted."[58] Second, the members would also have an incentive to cooperate in order to end the meetings early so they could spend time with their families. Third, with their spouses there, the Directors would also have less time to meet with one another and discuss any opposition they might have to Hock's proposals.

Hock also knew that most Directors tended to leave the meetings early if they felt that the final day's topics were unimportant, so he kept the eighty-percent authorization goal off the agenda. Hock had learned the time of every Director's return flight, and as those who would oppose the resolution drifted out, he adjusted his mental vote tally accordingly. After enough Directors had left, Hock switched topics and introduced the resolution, which was well-received by those remaining, and unanimously adopted. The larger issuers always wondered when they had ever voted for such a goal, but as the terminals were installed by most merchants, the reduction in their fraud losses convinced them that it was perhaps the right idea after all.[59]

8.3.5 Adoption and Consequences

Although the Board adopted the resolution, it actually took seven years to reach the eighty-percent authorization goal. Despite a set of economic incentives that enabled the terminals to pay for themselves within weeks or months, the merchants did not rush to install them. Bill Powar, who took over the POS business unit in 1982, explained that medium and large-scale merchants formulate and implement their technical plans according to a multi-year cycle, and will not consider the adoption of a new technology until the next planning stage. Although Visa members could demonstrate that the terminals would provide a net benefit for their merchants,

[57] Cleveland, p. 22.
[58] Ibid., p. 20.
[59] Ibid., pp. 22–23.

it often took several years before the merchants were ready to purchase and install the terminals in their stores.[60]

During this elongated adoption period, the terminal manufacturers continued to enhance their devices. As noted earlier, the first dial terminals performed authorization only; merchants were still required to complete and deposit paper sales drafts. By the mid-1980s, most terminals also supported data capture, meaning they could store the details of each transaction and transmit them electronically to the acquiring processor each night. These types of terminals also featured printers that could automatically dispense cardholder and merchant receipts, completely eliminating the need for the old multi-part paper sales drafts and card imprinters.[61]

When manufacturers began offering these data capture terminals, Visa USA again offered a special Terminal Interchange Reimbursement Fee (TIRF) for transactions that were not only authorized through the terminal, but also cleared electronically within three days. Like the original TIRF, this new rate provided enough incentive for acquirers and merchants to mass-adopt the data capture terminals, although the pace was again gated by technology planning and implementation cycles.[62]

As these terminals were installed, the paper largely disappeared from the system. Merchants still retain the paper receipts printed by the terminals, but they no longer deposit them with their acquiring bank. Just as with BASE II, the only time a merchant is required to produce the paper is during a dispute, making the exchange of paper the exception rather than the rule.[63]

However, the adoption of data capture terminals altered the system in a more profound and fundamental way than simply eliminating the paper. It transformed the way the banks viewed the acquiring business, which now resembled data processing far more than that it did banking. Banks had always justified their merchant discount fees based on the costs of supplying sales drafts, imprinters, training, and draft processing. The terminals reduced or eliminated most of these costs, and the TIRF resulted in further downward pressure on the discount fees, changing the business model into one completely based on scale. To survive, a merchant processor needed to acquire as many transactions as possible, and by the mid to late 1980s, banks found themselves either incapable or uninterested in providing the necessary computing infrastructure. As a result, many banks decided to outsource this function to more specialized firms such as First Data Resources (FDR), or exit the merchant side of the business entirely. Throughout the 1970s and early 1980s, nearly every bank was both an issuer an acquirer, but by the late 1980s, members tended to specialize in one function or another.

[60] Powar interview.
[61] Powar, Derman, and Fojtik interviews.
[62] Powar interview.
[63] Derman interview.

8.4 Conclusion

In this chapter, we reviewed how Visa fully automated the last, but perhaps most important, area of the payment system: the point of sale. They accomplished this by not only stimulating the development and adoption of inexpensive POS dial terminals, but also establishing a common standard for encoding the card. With the mass-adoption of data capture terminals in the mid-1980s, the Visa system was capable of processing transactions entirely in electronic form, a dream envisioned by Hock when he was forming NBI, and expressed in the BASE IV design.

However, I have yet to discuss perhaps the most important element in Hock's electronic value exchange system: access to deposits through the debit card. In Hock's world of EVE, cardholders should be able to access any pool of value they possessed, and his debit card was the key to accessing first deposits and ultimately investments. Although this seemed a natural step for Hock, the debit card was highly controversial for the larger banks, as they saw it as a threat to their own ATM and EFT plans. Visa USA also made a number of other moves during the late 1970s and early 1980s that caused the member banks to question whether Visa USA was now their competitor as well as their coordinator. It is to these events, and the ways in which they sparked debates about the roles Hock and the central Visa organization should play, that we now turn.

Chapter 9

Negotiating Roles: Debit Cards and Other Controversies

In the previous chapter, we saw how Visa fully automated the point of sale (POS) by establishing a common standard for encoding the card, stimulating the development of inexpensive POS dial terminals, and encouraging their widespread adoption. Once this occurred, Visa transactions could be processed entirely in electronic form, and the Visa system began to resemble the electronic value exchange system Hock had always desired.

In this chapter, I will complete my historical narrative by expanding upon something that was only hinted at in the previous chapter: conflicts over what role Hock and the central organization should play, both in the Visa payment system and the banking industry as a whole. The central Visa organization was created to facilitate and coordinate those activities that members could not perform alone, but we should not assume that the boundaries of these activities were ever "natural" or obvious to all those involved. Indeed, different groups in the system had different understandings of what the central organization should do, and more importantly, should not do.

When NBI formed in 1970, the members assumed that the organization's role was solely to coordinate the interchange of credit card transactions, that is, fix the existing BankAmericard system. As we have seen, Hock's conception of his and NBI's roles was far larger. Hock was intent on creating the premier system for electronic value exchange, and credit cards were only an initial piece of that plan. Thus, the actual role Hock and the central organization ended up playing had to be "worked out" as the system evolved. This working out was accomplished through various power struggles, some quiet and calm, others public and bitter.[1]

Although Hock and the members held different views about NBI's role, they were largely compatible for the first half of the 1970s. The establishment of the bylaws and operating

[1] The idea that different actors in a network attempt to "enroll" one another comes from Actor-Network Theory. See Callon, 'The Sociology of an Actor-Network'; Latour, *Science in Action*; Latour, *Reassembling the Social*; and Law.

regulations, the development of BASE I and II, and the formation of IBANCO, could all be seen both as actions to fix the BankAmericard system *and* necessary building blocks for a larger electronic value exchange system. Although both Hock and the members saw these developments as important, they did so for different reasons.

However, all of this changed in the mid-1970s when Hock proposed what would become NBI's first signature-based debit card. For Hock, this was the next logical step in his electronic value exchange plan, but for the members, it was an unwelcome expansion of the role they had assigned the central organization. In 1979, Hock again challenged this more limited role by signing a direct merchant agreement with JC Penney, one of the three largest US retail chains at the time, bypassing a merchant acquiring bank. As a result of these controversies, the members began to accuse Hock of empire-building, and acting more like their competitor than their coordinator. As Hock and the members continued to struggle over their respective roles, tensions increased, Hock's power of persuasion began to fade, and he was ultimately forced out of the organization in 1984.

9.1 Entrée and Early Debit Cards

The topic of Visa as a primary cause of mounting consumer debt is a perennial favorite of the popular press, and even a few academics.[2] These authors typically accuse Visa of encouraging consumers to abandon thrift, finance extravagant lifestyles, and accumulate larger and larger amounts of burdensome or even unmanageable debt. While this kind of accusation could be made against some of the larger American issuers, it is difficult to maintain against Hock and the central Visa organization itself. Recall that Hock's goal was not to automate the issuance and use of consumer credit *per se*, but to enable the exchange of any kind of value the consumer might possess. There is little about the Visa payment system, both technically and organizationally, that assumes the value being exchanged comes from a line of credit. These authors rarely mention that the banks in France, New Zealand, and Japan have always used cardholder deposits to settle their Visa transactions. They also tend to ignore that since 2003, the Visa system has processed more debit than credit transactions each year, both in terms of number and value, and the ratio of debit to credit transactions continues to rise.[3] These authors also seem to be unaware that Hock and his staff began researching and developing the first international debit card in 1973, just three years after the organization formed.[4]

In this section, I will present Hock's concept for that card (which was fairly radical at the

[2] For example, see Manning.

[3] 'Visa Global Debit Card Volume Surpasses Credit', `http://corporate.visa.com/md/nr/press217.jsp` (accessed on 18 April 2007). See also, W Lee, 'Visa Debit Card Volume Tops Credit Worldwide', *American Banker* (21 April 2004), p. 11. In the US, the number of debit transactions exceeded credit a few years earlier. When Manning's book was published in 2000, the debit transactions had already grown to a quarter of the total US dollar volume and were growing more rapidly than credit transactions. See Jennifer Kingson Bloom, 'Visa Card Use Snowballing', *American Banker* (13 March 2000), p. 14.

[4] 'Visa Debit Card Services' (1977), p. 1.

time), discuss NBI's product based on that concept, why it would appeal to consumers, and why most of the membership initially rejected it. This story highlights that Hock was playing a different role than most of the members assumed, and their rejection of the card can also be seen as an attempt to rein in Hock's larger ambitions.

9.1.1 The Asset Card Concept

In 1973, after BASE I was put into production and while BASE II was being developed, Hock began speaking publicly about the need for a new card-based payment service that could access not only a line of credit, but also deposits and investments.[5] He referred to this as an "asset card" to distinguish it from the more limited notion of a credit card:

> In the context of customer use, "credit card" has always been a misnomer. Certainly from the user viewpoint, a bank card is solely used to create debits. That is, to dispose of value owned by or to be earned by the user. The term "credit card" is... a classic example of naming and marketing the product from the perspective of the supplier rather than the user, thus making it unacceptable to a large number of prospective customers.[6]

Hock understood that the names we assign to things tend to influence the way we think about them. This is why the new name for his organization ("Visa") had to be free of any geographic ("American"), service ("Credit or Charge"), or form ("Card") connotations. Similarly, he wanted the bankers to speak about "asset cards" instead of "credit cards" because the latter tended not only to repel those averse to the idea of buying things on credit, but also to limit the ways in which the bankers thought about what the card could do.

Hock argued that the term "credit card" had conditioned bankers to think of bankcards only as vehicles for consumer credit, and not as generalized payment devices that could access any kind of asset the cardholder might possess. Limiting the card to credit lines actually made no sense to Hock, as he saw very little difference between credit and deposit accounts:

> Bank card accounts, savings accounts and checking accounts have only two substantive differences: First, the time at which the bank requires the customer to make credits—deposits or payments—to balance the ledger, and second, the party to receive interest and at what rate depending on the balance struck.[7]

In other words, credit and deposit accounts are actually two different sides of the same phenomenon: an account. That account can contain funds loaned to the customer by the bank, or funds loaned to the bank by the customer. In fact, a single account can oscillate between the two states. If you overpay your credit account, the bank will happily hold on to the money,

[5] Information on the asset card concept comes from interviews with Honey, Derman, and Russell, as well published sources cited later.

[6] Dee Hock, 'Bank Card Industry Must Press for Legislation on National Level', *American Banker* (3 October 1973), p. 4.

[7] Ibid.

express your balance as positive, and apply those credits to your next statement. If you overdraw your current account, the bank can extend you temporary credit, express your balance as negative, and apply your next deposit as repayment. Hock argued that this separation between deposits and credit lines was more of a historical accident than a natural distinction, and there was no need to perpetuate it as banks moved into the world of electronic value exchange.

If there is little difference between credit and deposit accounts, Hock continued, then the mechanisms for exchanging the value contained in those accounts must also be essentially the same:

> Checks, like bank cards, are simply mechanical devices for exchanging value by debiting accounts. It is only by custom that credits are required in advance, that third party acceptance carries no assurance of payment and that MICR [Magnetic Ink Character Recognition] encoding on paper governs the mechanics of clearings. In short, that checking accounts are designed more for the convenience of banks in acquiring funds than for customer convenience in the market place.[8]

Hock proclaimed that his asset card could do everything a cheque could do, and do it *better*, at least from the perspective of consumers, who were quickly becoming the most important customer segment for the future of banking. The cheques issued to the increasingly mobile American consumer were unhelpful, as they were not typically accepted outside the issuer's local area, primarily because they were not guaranteed. The BankAmericard had demonstrated that transactions could be guaranteed on a national, and even international scale, and the asset card would be built upon the same organizational and technical infrastructure.

Furthermore, because bankers made an artificial distinction between deposit and credit accounts, their cheques were unnecessarily inflexible. Cheques required customers to make deposits in advance of their needs, but to do so, customers had to visit a branch—which Hock noted was inaccessible eighty-two percent of the time—and stand in time-consuming lines.[9] Why, Hock asked, were banks willing to extend their customers ample credit on their BankAmericards, yet not connect that credit to their customers' demand deposit accounts to cover overdrafts? He suggested:

> Bank card accounts could quickly be converted to banking accounts by revising software to carry either credit or debit balances, and marketing the service as a value exchange device with an optional credit feature.[10]

Not surprisingly, this new service would be accessed by Hock's asset card, which would enjoy the same nationwide (and eventually worldwide) acceptance as the BankAmericard. This was the essence of electronic value exchange, and by implication, the future of banking.

[8] Hock, 'Bank Card Industry Must Press', p. 4.
[9] Ibid.
[10] Ibid.

9.1.2 NBI's Asset Card

The task of turning this concept into an NBI product was given to Tom Honey, the product developer already introduced in Chapter 7. Although Honey could not control how the banks managed the relationship between their deposit and credit accounts, he could easily design and market a new kind of card that would primarily access funds in a demand deposit account rather than a credit line.

He designed the new card to be a companion to the existing BankAmericard. From an operational perspective, it would look and function almost identically to its elder: it would feature the same blue, white, and gold bands design; would identify the cardholder to the merchant and the system; and would be accepted at any location that currently honored the BankAmericard. When used for purchases, merchants would process the transaction in exactly the same manner: they would call for authorization if necessary; complete a paper sales draft; and deposit it at their acquiring bank. A draft produced from this card would also clear and settle electronically through BASE II just like a credit sales draft. However, once it reached the issuer, instead of adding the draft to an existing credit account, the issuer would handle it more like a cheque, directly debiting the cardholder's deposit account. No further billing or collection would be necessary.[11]

The asset card was similar in many ways to an electronic cheque, but with four important differences. First, the asset card separated account identity from the payment instrument, allowing the card to be used in multiple contexts and devices. Banks were already issuing cards for use in their ATMs, and Hock wanted to replace those special-purpose cards with the more general-purpose asset card, which could also be used at any BankAmericard merchant. He was also thinking ahead to his BASE IV world, where the card could be used in electronic POS terminals, or any transaction origination device.

The second, and perhaps most critical difference was that asset card transactions would be *guaranteed*. In the cheque payment system, issuing banks could simply return a cheque to the merchant if the account had insufficient funds, but asset card transactions would be subject to the same payment rules as BankAmericard transactions. If the merchant followed the authorization rules, the merchant was guaranteed payment, even if the source account did not have enough funds. Because floor limits were still in place at this time, overdrafts were thus a distinct possibility, implying that issuers would sometimes be forced to extend temporary credit to cardholders.

The third difference was the scope of acceptance. Because asset card transactions were guaranteed, merchants could accept cards issued from any bank, even if it was outside the local area. In fact, Hock's principle of "universality," which was canonized in the bylaws, operating regulations, and merchant contracts, specifically required merchants to accept *any* card with the blue-white-and-gold bands design, regardless of who issued it or how that organization settled

[11] Honey interview.

with the cardholder. While cheques eliminated the need to cary large amounts of cash, they were essentially useless outside the issuer's area or at stores where the customer was unknown and thus untrusted. The asset card promised to blend the best of cash and cheques: it would be accepted almost as readily as cash, yet be as convenient and safe as a cheque.[12]

The last critical difference was in the pricing of the service. As discussed in Chapter 3, one of the goals of the new Federal Reserve System was to eliminate the practice of discounting cheques cleared through interchange. By the early 1970s, the Fed and State legislatures had enforced par clearing between all banks, and merchants had come to expect full payment for the cheques they deposited. Although NBI considered the asset card as a replacement for cheques, it was priced in the same way as existing BankAmericard transactions: merchants paid a two to six percent discount fee to acquirers, and acquirers paid the standard interchange reimbursement fee to the issuers. These fees offered a significant benefit to acquirers and issuers, but merchants objected, arguing that they had signed up to accept a credit card, not a discounted replacement for cheques.

NBI justified this pricing by appealing to the guaranteed payment feature, and noting that they did not benefit from the same government subsidies enjoyed by the Federal Reserve. However, a few merchants considered it an attempt to reintroduce a non-par cheque, and refused to accept the asset card when it was first issued.[13] Unfortunately for them, their contracts specifically required them to accept all cards bearing the blue, white, and gold bands design, and most merchants eventually capitulated, primarily because volumes remained quite low during the 1970s.[14]

9.1.3 “Credit Paranoia”

Based on his experience at Wells Fargo, as well as prior consumer research done by NBI, Honey surmised that the asset card would be most attractive to a segment of the population that he described as having “credit paranoia.” This group was not opposed to payment cards *per se*, but was reticent to adopt a card that accessed a line of *revolving credit*.[15]

Consumer credit has always been a rather emotional subject in the US and the public discourse surrounding it has often been at odds with the actual practices of Americans. Nocera explained this well:

> Consumer credit. . . has always occupied a peculiar place in the American psyche. On the one hand, there is no aspect of personal finance more likely to inspire

[12] This is not to say that cheques are somehow less prone to fraud than credit card transactions. The point was that asset card transactions should be no more dangerous than a cheque.

[13] Jeffrey Kutler, ‘Phoenix Merchant Fights Visa Debit Fee’, *American Banker* (9 June 1977), p. 1.

[14] By 1996, the ratio of debit to cheque purchases had risen to such a level that several merchants, led by Wal-Mart, sued Visa and MasterCard, demanding the right to accept the credit but not the debit card. Visa and MasterCard eventually settled, allowing merchants to accept one type of card without having to accept the other. They also established different interchange reimbursement fees for the two types of cards, allowing acquirers to charge a lower discount for debit than for credit. See Evans and Schmalensee, pp. 291–294.

[15] Honey interview.

> anxiety and even fear. At any moment in our history, one can find ringing denunciations of consumer credit and “usurious” interest rates, calls for reform, worries that things have finally gotten out of hand. “Rather go to Bed supperless than rise in Debt,” wrote Ben Franklin, and Americans have been echoing that sentiment ever since. . . .
>
> On the other hand. . . despite the denunciation, despite the free-floating anxiety, Americans have always borrowed money to buy things—if not from a bank, then from *somebody*. . . . There isn't another Western country that has relied so heavily on consumer credit; between 1958 and 1990, there was never a year when the amount of outstanding consumer debt wasn't higher than it had been the year before.[16]

Bank-issued credit cards, which made consumer credit even more widely available, met with a similar response. Despite vituperative condemnation from political officials and stinging articles in the popular press, many Americans still chose to adopt them. More than 38.7 million Americans carried at least one bank-issued credit card by 1974 and that number was steadily increasing.[17]

However, Americans were often conflicted about the proper use of bank-issued credit cards. Professional organizations, most notably the American Medical Association (AMA), passed resolutions in the late 1960s prohibiting their members from accepting credit cards for their services.[18] Although the AMA did eventually allow doctors to accept credit cards in 1971, they were still forbidden to advertise the fact outside of their offices.[19] Several states, such as New York, also passed laws prohibiting medical professionals from accepting credit cards for payment. There was just something unseemly about allowing consumers to finance their health care on credit.

There were also categories of purchases for which Americans in the 1970s would never dream of using a credit card. Cultural norms dictated that certain types of goods and services could be purchased with a credit card, but anything else should be budgeted for, and purchased with cash.[20] Today we commonly use credit cards to buy basic necessities such as groceries in order to earn air miles or some other type of rebate, but thirty years ago, buying these types of items on credit was a sign of financial desperation. It communicated something tragic about one's social position, and was thus something to be avoided.[21]

In previous research, NBI had determined that about half of all Americans who had demand deposit accounts did not carry any kind of bank-issued credit card.[22] When asked why,

[16] Nocera, p. 20.

[17] Hock, ‘Bank Cards’, p. 18. For a classic critical article from the popular press, see O'Neil, pp. 48–50.

[18] Stallwitz, p. 56.

[19] ‘Card May Be Solution to Unpaid Doctor's Bills’, *American Banker* (26 January 1971), p. 1.

[20] H Lee Mathews and John W Slucum, Jr, ‘Correlatives of Commercial Bank Credit Card Use’, *Journal of Bank Research* 2:4 (1972), pp. 21–27.

[21] This is similar to Zelizer's observation that we often treat money earned from different pursuits in different ways. See Zelizer, p. 3.

[22] ‘Visa Debit Card Service: A Digest of Key Research Findings’ (March 1977).

the most common response by far was that credit cards tempted one to overspend and incur unmanageable debt. This group was unlikely to ever adopt a credit card, but a card that settled against their deposit account might be attractive.

Honey expected that the asset card would appeal strongly to those concerned about overspending, and would ultimately been seen as acceptable for any purchase currently made with cash or cheques. To test the idea, he hired the Field Research Corporation to conduct 1,675 in-depth interviews in various US metropolitan areas across the nation.[23] The results confirmed many of his suspicions, but also revealed some interesting surprises.

First, over sixty percent of those surveyed showed some interest in the asset card. Interest came both from those who had a credit card and those who did not. Those who paid their bill in full each month were interested because it would eliminate the bill, but surprisingly, the strongest interest came from those who financed at least some of their purchases. When asked why, these cardholders responded that they saw it as a convenient replacement for cash and cheques, not as a replacement for their credit card, indicating that they wanted to carry both and choose when to pay with deposits, and when to pay with credit.[24]

Second, those interested in the card were actually willing to pay for the privilege of using it. Honey had expected this given his experience helping to develop the Wells Fargo Gold Account, but the member bankers in the advisory group were shocked.[25] They had always assumed that payment services such as cheques had to be given away free in order to entice deposits. A cardholder fee combined with the merchant discount or interchange fee would likely make the service at least self-supporting, if not outright profitable.

9.1.4 Member Reactions

Although consumers seemed to like the asset card concept, Honey encountered significant resistance when he began pitching it to many of the member banks. This may seem surprising as the asset card itself had a rather attractive value proposition: it was a fairly simple extension of the BankAmericard system; would reintroduce the revenue stream banks used to receive from non-par cheques; and might even generate new monthly cardholder fees. However, many of the card managers worried that the asset card might "cannibalize" their existing BankAmericard programs, many of which were just starting to become profitable. Honey tried to use the consumer research to show that cardholders would use the asset card as a companion to their BankAmericard, and not as a replacement, but most members remained unconvinced.

But there was another important and deeper reason why these banks resisted the asset card. Many of these banks were not against the idea of a debit card *per se*, but they were against the idea of a debit card controlled by NBI. They saw the asset card as a challenge to their own

[23] 'Visa Debit: A Digest'; 'Visa Debit Card Services'.

[24] Ibid.

[25] The Gold Account offered cheques as well as other consumer banking services for a flat monthly fee (Honey interview).

ATM and Electronic Funds Transfer plans, most of which had different assumptions about the required technological infrastructure, and how it should be shared.

Electronic Funds Transfer (EFT) was a rather unavoidable subject for US bankers in the 1970s, and by 1974, most of the NBI member banks were pursuing some kind of EFT plan. These plans ranged from the rather simple installation of a few ATMs in each branch, to more complex regional networks of ATMs, cheque guarantee programs, and even point-of-sale debit card systems. Although most of these systems faced the same kind of operational issues the credit card programs had recently overcome, these banks did not consider the card associations to be the proper foundation upon which to implement their EFT plans. This was due at least in part to a particular social dynamic that was operating in most US banks at this time.

Credit Cards vs. "Real" Banking

Several interview sources commented that during the 1970s and early 1980s, there was a deep cultural and political divide between the credit card operations and what they called the "deposit side" of the bank. Hock lamented in a 1974 speech that "In the minds of the executive management of most banks, bank cards are little understood and rarely thought of as a real part of the bank."[26] Tom Cleveland, Visa's treasurer and CFO, also pointed out this attitude:

> In bad economic times, the Visa portfolio could deliver 70–80% of a bank's bottom-line profits, 10–15% was normal, but the card operations were never given the respect they deserved. They were not viewed as banking. They were the new stepchildren in the family of banking that everyone wished would stay quiet and unseen, but would keep working harder than ever.[27]

While the directors of these banks often recognized the potential profits to be made from the credit card, the deposit side of the bank still considered it a somewhat questionable venture: something that might be useful in attracting new consumer deposits, but not something the bank should ever focus upon or integrate as a core function. For them, the credit card was just a new form of unsecured consumer lending, which was an activity more commonly associated with disreputable finance companies, corner pawn shops, and loan sharks. The high fraud losses resulting from the unsolicited mass-issuance of cards in the late 1960s, only further reinforced this view.

For these bankers, *real* banking consisted of two primary activities: taking in more and more deposits, so that you could make larger and larger commercial loans. This is how banks traditionally made their profits, and as such, those involved with deposits and commercial lending tended to have more political power within the bank than those involved with consumer lending and credit cards.

[26] 'Banking Lacks National Structures to Deal Adequately with Change', *American Banker* (25 September 1974), p. 1.

[27] Cleveland, p. 21. Of course, this was not completely universal. A few progressive or innovative banks did view the cards as critical to their future, and these were the banks that either accepted the asset card, or were pursuing their own advanced EFT solutions. However, these banks were the exceptions, not the rule.

This power differential often took a physical form as well. The card operations center was commonly housed in the least desirable location of the bank building, often in the basement or attic. This was partly due to the noisy, finicky computing and sorting equipment, but card managers also felt an unspoken desire to keep them far away from the stately lobbies and offices where the "real" banking took place. Don Jutilla, the BankAmericard program manager for Puget Sound National Bank recalled his department's accommodations:

> [the bank] never really thought we'd make it. They started out by putting us on the seventeenth floor of a sixteen-story building. We had a gold spire on the top of the building... they put us up in there, thinking that we were probably just going to fade away and they wouldn't have to build a building for us or anything.[28]

Differing Views of EFT

Because of this cultural and political divide within the banks, it was typically those on the deposit side who formulated and controlled the bank's EFT plans, not those on the credit card side. Instead of rethinking their assumptions and practices in light of the new technology and consumer needs, the deposit bankers tended to develop their EFT plans as extensions of their current practices, which were quite different from those of the credit card associations. As a result, their plans tended to differ from the asset card in at least three important ways.

First, bankers on the deposit side were typically much more conservative and cautious, so many of their plans opted for an incremental approach towards EFT involving ATMs and cheque authorization or guarantee systems.[29] Instead of replacing cash and cheques with a debit card, these banks sought to make it easier for customers to obtain cash, and safer for merchants to accept cheques. Any card they issued merely served as an additional form of identification, or the mechanism for accessing a cheque authorization service, which simply compared the account against a negative list. Some banks did go a step further by providing special terminals that transformed the cheque into a guaranteed payment instrument, but only in particular stores within a limited area.[30]

Second, for those that did plan to offer a point of sale debit card, they often required a different technical infrastructure than what NBI was proposing for the asset card. This is best explained by realizing that these bankers tended to see the point of sale as a special kind of ATM, while NBI typically viewed an ATM as a special kind of merchant. As a result, these

[28] Jutilla interview. Jutilla noted that because his bank did embrace consumer lending more than others, he was actually treated better than most.

[29] For example, see 'Nashville Banks Cautious on EFT', *American Banker* (11 December 1975), p. 1.

[30] 'Security Pacific, California, Contracts for POS System for Merchants', *American Banker* (7 October 1976), p. 2. Interestingly, the Nashville banks did not offer a guarantee because they were concerned that doing so would turn the merchant into a potentially illegal "branch," as cheques were not guaranteed by law until they were presented to, and accepted by, the issuing bank. Whether an ATM or a POS terminal constituted a "branch" subject to state branching laws was also disputed in the courts. See *Independent Bankers Association of America v. Smith, Comptroller of the Currency of the United States*, No. 75-1786, United States Court of Appeals For the District of Columbia Circuit 175, decided 23 March 1976.

bankers wanted a similar operational environment at the point of sale as they had for ATMs: electronic terminals; supporting the entry of a Personal Identification Number (PIN) for cardholder authentication; authorizing and clearing every transaction at the time of sale with a single message. In contrast, NBI's asset card was designed to operate in the same environment as the BankAmericard: accepted worldwide; at merchants with or without terminals; authenticated only by a signature; authorized only when the transaction exceeds the merchant's floor limit; and cleared at a later time with a separate message. It would take many years for merchants to adopt the technical infrastructure necessary to support an ATM-like transaction at the point of sale, but these banks felt it was necessary to avoid excessive fraud and overdrafts.

Third, their plans often embodied a different model of competition and cooperation than NBI's. As already noted, the deposit side of these banks viewed the point of sale as a special kind of ATM, and they also viewed the ATM as a special kind of branch. Branches provided a competitive presence in a service area, and as such, were not "shared" with rival banks. Similarly, ATMs were not typically shared with rivals because they too provided a competitive presence. EFT devices and services at the point of sale were treated the same way; these banks were often willing to cooperate with their correspondent banks far away, but wanted a way to deny access to their competitors down the street. NBI's principles of universality and open membership would have made that impossible, so these banks formed new associations with different rules, or simply kept their technology proprietary, licensing it to selected correspondent banks.

This was quite common in the regional ATM networks that would eventually consolidate into the PLUS, Cirrus, and STAR networks, but one example in the world of debit cards was Worthen Bank and Trust's "Moneycard" service, which it began licensing in 1975.[31] Their Chairman and CEO described their strategy:

> Worthen believes that the best avenue for development of electronic funds transfer technology is a joint approach involving regional money center banks, such as Worthen, in harmonious cooperation with correspondent banks located in towns and cities of all sizes throughout the state.[32]

Despite rhetoric such as "joint approach" and "harmonious cooperation," the operative phrase in this quote is "correspondent banks." Worthen was willing to share their debit card service with other banks, but only those with which they had a non-competitive relationship. They also wanted to retain complete ownership over the technology and marks, merely licensing them to their correspondents much like the BofA licensed the BankAmericard prior to NBI.

In many ways, NBI's asset card was a collision of social worlds. It looked and acted like a credit card, but it directly accessed consumer deposits. It was controlled by a card association,

[31] For a review of ATM network sharing policies up until 1984, see Steven D Felgran, 'Shared ATM Networks: Market Structure and Public Policy', *New England Economic Review* (January/February 1984), p. 27.

[32] Edward M Penick, quoted in 'Worthen B&T Offers Debit Card Service', *American Banker* (12 September 1975), p. 1. Worthen is the same bank that challenged NBI's dual membership ban. See Section 7.1.2 beginning on page 123.

and promoted internally by the credit card department, and as such, threatened to upset the balance of power between the consumer credit and deposit sides of the bank. It also brought with it a certain operating environment that was radically different to that of cheques and ATMs, and many of the banks were uncomfortable with the possibility of increased or uncontrollable overdrafts. It brought with it a particular set of rules that required the banks to cooperate with their direct competitors. But perhaps the worst part was that it would allow Dee Hock to dictate how the banks would enable electronic access to their deposits. Ken Larkin, the BofA executive who managed the BankAmericard licensing program before NBI, put it bluntly:

> As long as [Hock] kept to credit cards, banks were willing to give him leeway. The moment you get into the debit card, you're talking about more than $1 trillion [in deposits]. You're hitting bankers where they live. They weren't going to let anybody, *especially* Dee Hock, tell them how to manage their deposit stream.[33]

9.1.5 Enter Entrée

Although the number of potential issuers looked to be small, NBI announced in August 1975 that its new asset card, now named Entrée, would be available to issuers starting in October.[34] Initially, Hock wanted the asset card to look exactly the same as the BankAmericard, as he considered the manner in which cardholders settle with their banks to be a matter of privacy. However, since electronic POS terminals with credit/debit options were not yet widely available, cardholders needed a way to distinguish between the two types of cards, so the separate name Entrée was used. Unfortunately for NBI, the separate name also had the side effect of providing merchants with an easy way to identify, and potentially refuse, what they considered to be the reintroduction of a non-par cheque. When NBI adopted the name Visa in 1977, they dropped the name Entrée from the asset card and let issuers put their own name in the blue band, making it much more difficult to distinguish from the credit card.

When the service officially launched in October 1975, about fifteen banks promised to issue the Entrée card, but the rest remained uncommitted.[35] One of the first to issue was City National Bank and Trust of Columbus, Ohio, the same bank who tested a similar type of card back in 1971.[36] They were soon joined by Pittsburgh National Bank, as well as banks in Colorado, Arizona, and Louisiana.[37] As Honey had predicted, these early adopters found that the Entrée card did not cannibalize their credit card volumes, and at least one bank was able

[33] Ken Larkin quoted in Nocera, p. 308.

[34] Geoff Brouillette, 'NBI Set to Launch Debit Card Plan as Early as October', *American Banker* (28 April 1975), p. 1; Phillip Brooke, 'New NBI Debit Card is Named ENTREE', *American Banker* (22 August 1975), p. 1. Confusingly, after the asset card was renamed Visa Debit, the name Entrée was reused for a separate joint venture between Visa/PLUS and MasterCard/Cirrus in the 1980s (Derman and Harrison interviews).

[35] 'Visa Debit: A Digest', p. 1; 'NBI Announces Entree Card', *Payment Systems Newsletter* (September 1975), p. 1.

[36] See page 85.

[37] 'Pittsburgh NB Will Offer Entree Card', *American Banker* (3 March 1976), p. 6. '25 Colorado Banks Will Start Offering Entree', *American Banker* (23 April 1976), p. 2.

to charge a monthly cardholder fee.[38] They also discovered that their Entrée cards were used far more often than their BankAmericards, yet they experienced less fraud, primarily because cardholders were more careful with what they regarded as their own money.[39]

The number of issuers grew slowly throughout the 1970s, reaching only 90 by the end of the decade, causing the press to comment on the "disappointing acceptance of the concept over the past five years."[40] The adoption rate began to increase at the start of the 1980s, growing to 306 by the end of 1981. The widespread adoption of POS terminals helped matters a bit, as most transactions could now be authorized, and overdrawn accounts could be shut off quickly.[41] However, throughout the 1980s, most banks continued to issue ATM and debit cards that carried one of the marks from the other PIN-based, single-message networks, such as PLUS, Cirrus, or Interlink, and not Visa.[42]

In 1993, Visa began a major effort to convince the top fifty US banks to issue their signature-based debit card, pointing out that it would generate far larger interchange revenues than those offered by the PIN-based ATM and debit networks. They also renamed the card to "Visa Check" to further emphasize its ability to replace cheques at the point of sale, and distance it from the word "debit," which many consumers associated too closely with "debt." A large national advertising campaign cemented the concept, and when American supermarkets began accepting Visa cards, their efforts began to pay off: the number of transaction per year on Visa USA's debit cards rose from 87 million in 1990 to 6.5 billion by 2002.[43] As already noted, Visa now processes more debit than credit transaction each year, both in terms of number and dollar volume, in the US and worldwide.

9.1.6 Coordinator or Competitor?

Since the asset card was ultimately adopted by the members, it is tempting to think of it as simply a product ahead of its time. That may have been true to some extent, but it is also helpful to view it as the locus of a power struggle between the central Visa organization and its members over what it should and should not be doing. From Hock's perspective, the asset card was simply the next logical step in creating his system for electronic value exchange, but from the membership's perspective, the asset card was a break from NBI's past activities, and an attempt to control their EFT plans. In their minds, NBI existed to coordinate their credit

[38] 'Visa Debit: A Digest', p. 2. Louisiana National charged one dollar a month after the first six months, which was just enough to weed out the inactive cardholders.

[39] Honey interview.

[40] 'NCNB To Become Biggest Bank in Visa Debit Card Program', *American Banker* (8 March 1979), p. 3.

[41] Jeffrey Kutler, 'Age of the Debit Card is Coming', *American Banker* (9 March 1981), p. 1. 'Debit Card Volume Rises 187% At Visa', *American Banker* (14 December 1981), p. 9.

[42] Only eleven percent of US ATM cards carried a Visa mark by 1994 (Evans and Schmalensee, p. 206). Honey believes that Hock mistakenly ignored the growing threat posed by these networks, favoring instead to focus on building his international organization (discussed in section 9.3.1). Visa eventually had to acquire PLUS and Interlink in order to gain a foothold in the single-message, PIN-based transaction environment.

[43] Ibid., pp. 206–211.

card programs, not compete with their EFT offerings.

To be clear, Hock did not wish to control the members' EFT plans as much as he wanted a chance to offer consumers what he thought was a truly better option. He was concerned that member banks would repeat the early history of bank credit cards, either by waiting for the ideal solution while losing their customers, or racing to issue ill-conceived, localized debit cards with even less geographical range than a cheque. However, the banks often did not see the need for national acceptance, and were more interested in reducing the costs of handling cheques than in improving them.[44] Both Hock and Honey continued to push the member banks to issue the asset card, but the lack of support from the issuers during this time period ultimately kept it out of the hands of consumers.

By resisting the asset card, the member banks attempted to establish limits on the role they thought Hock and the central organization should be playing. Of course, Hock continued to challenge these limits, attempting to expand the activities his organization coordinated, and redefining his own role as he saw fit. The tension this created between himself and the membership can be observed in the frustration he began to express in his annual conference speeches. He accused the bankers of being "highly suspicious of new structures... historically successful at reaction rather than action... [and] comfortable with the theology of centralized noncompetitive clearing of checks." He warned them that their stubborn adherence to existing structures and practices might soon turn them into an anachronism:

> Certainly the vast religion of electronic value exchange ought to be able to worship effectively in more than the temple of demand deposits, checks, Federal Reserve clearings and interest on loans, and without the priests requiring that all belong to the same sect. But I suppose that depends on whether it is the form or the substance of banking that we worship most. The congregation we call customers may remain for a time, since the temple is familiar, but other preachers with powerful messages are talking just outside.[45]

9.2 The JC Penney Deal

The second controversy to discuss in this chapter is Visa USA's infamous 1979 deal with JC Penney, one of the three largest American retail store chains at that time. This episode, more than any other, sparked extensive debates about what Visa's role should be, and whether Visa was now acting more like a competitor than a coordinator.

To understand this deal and the controversy it created, we must first review the context in which it took place. In the late 1970s, very few national retail chain stores accepted bank-issued credit cards. K-Mart and Fred Meyer did, and Macy's accepted American Express to

[44] Bjorndahl interview.

[45] 'Bank Cards Are Real Banking, Too', *American Banker* (27 September 1974), p. 4.

attract tourists, but the largest stores—Sears, Montgomery Wards, and JC Penney—did not.[46] These stores had been offering their own, private charge or credit cards for many years, and by 1978, Sears alone had more cards in circulation than Visa and Interbank combined.[47] There was little economic incentive for these stores to accept bankcards: their own cards incurred no discount fees, and those that allowed customers to finance purchases generated additional interest revenue. Even if these revenues did not offset the costs of running their programs, they still considered them valuable, as they thought the cards engendered a sense of loyalty in their customers, one that they were reticent to give up by accepting bankcards.

The number of transactions these retailers generated was enormous compared to most of Visa's merchants, and Hock wanted Visa to process as many of them as possible. So far, the top three retailers had remained unified in their opposition to bankcards, but Hock also knew that large retailers are typically willing to break ranks if doing so can get them ahead of their competition. Sears was already the largest in terms of both sales and credit accounts, so Hock focussed his efforts on one of their close competitors: JC Penney.[48] Penney had 2,000 stores across the nation, and 13.6 million cardholders. On average, 42 percent of their sales were done on their private credit card, comprising a total volume of $4.5 billion in 1978. Penney's credit volume alone was over 20 percent of Visa USA's entire volume for the same period.[49] Capturing even a portion of that would result in a significant surge for Visa.

9.2.1 The Deal

After extensive negotiations with Hock, Penney agreed to accept Visa cards starting in the fall of 1979. However, instead of signing an agreement with one of Visa's member banks, which all other merchants had done, Penney signed the agreement *directly with Visa USA*.[50] Thus, Penney became their own merchant acquirer, and would pay fees to Visa USA only.

Although the lack of merchant acquirer revenues was bad enough, the agreement also resulted in lower revenues for the issuers as well. To close the deal, Visa USA developed a special electronic interchange reimburse fee (EIRF) rate that applied to any transaction authorized electronically and submitted to BASE II within three days. This was how Penney intended to operate, and the EIRF effectively gave them a twenty-percent discount on the interchange fees they paid to issuers.[51]

To make matters even worse, Hock arranged the terms of the contract without first consulting the Board of Directors. He claimed that Penney required this for secrecy, but he must

[46] Jeffrey Kutler, 'Penney to Honor Visa; 1st Big Retailer to Accept Bank Cards', *American Banker* (5 April 1979), p. 1.

[47] 'Sears Has More Cardholders Than Either Visa or Master Charge', *The Nilson Report* 182, p. 1.

[48] Various news reports listed Penney as either the second or the third-largest retailer at that time.

[49] Kutler, 'Penney to Honor Visa'. Visa USA's volume was $22.1 billion in 1978.

[50] 'Usurpation Feared in Visa, Penney Tie', *American Banker* (11 April 1979), p. 1.

[51] Jeffrey Kutler, 'Visa Sued on Price Fixing, Penney Tie', *American Banker* (27 June 1979), p. 1. This rate became the prototype for the two TIRF rates discussed in Chapter 8.

have also known that the Board was unlikely to agree to such terms in advance. At the end of a two-day Board meeting, Hock announced that there was one more item of business, and happily informed the Directors that he had convinced JC Penney to accept Visa cards.[52]

9.2.2 The Reactions

As one might imagine, when the Directors learned the *terms* of this deal, they did not react well. Although the inclusion of Penney would no doubt result in higher volumes, and ultimately lead to acceptance at the other national retailers, they all felt that Visa had cut them out of a very important and lucrative merchant contract.

Hock tried to justify the lack of a merchant discount by arguing that Penney did not need a merchant bank. They captured transactions electronically using their own equipment, authorized them directly with BASE I, and electronically submitted them for clearing to BASE II. All the services that a merchant bank typically provided were unnecessary for Penney, so why should they pay a merchant discount fee? Similarly, the EIRF was justified on the basis that Penney would be authorizing every transaction, and the three-day clearing meant that issuers would see transactions much sooner, allowing them to detect and stem fraud more quickly.[53]

After a lengthy debate, the board reluctantly agreed to approve the contract. Chuck Russell described how it was finally settled:

> The deal was the subject of great controversy amongst the directors until Ken Larkin spoke up and said something like: "Look my fellow directors, none of you could have signed Penneys, as evidenced by the fact that none of you *have* signed Penneys. Let's take our victory, and enjoy the fruits of Dee's labor." The board finally calmed down, but contemporaneously adopted a resolution forbidding the signing of merchants by Visa management or staff.[54]

This resolution was of course another move to limit the scope of Hock's role, and firmly establish which activities should be done by the central organization, and which should be left to the member banks.

When news of the deal reached the general membership, they were livid. One officer of a regional merchant bank responded, "I don't like it. I intend to raise hell about it when I get the chance."[55] He asserted that Penney should be associated with a merchant bank, if only arbitrarily, so as "to preserve Visa's traditional standing as an association rather than a competitor with some of its more prominent members." In other words, this deal made it look like Visa was now becoming their competitor instead of their coordinator.

National Bankcard Corporation (NaBANCO), a large merchant processor in Florida, went so far as to file an antitrust lawsuit against Visa USA, part of which sought an immediate

[52] Nocera, p. 306. Confirmed in Russell interview.
[53] Streeter.
[54] Russell correspondence.
[55] 'Usurpation Feared'.

injunction of the Penney arrangement.[56] The NaBANCO suit is better known for its challenge of the interchange fee in general, but NaBANCO was also upset that Visa had denied them revenues they could have earned by acquiring Penney's transactions. They questioned whether Visa, as an association of competitors, had the right to compete with its members.

John Reynolds, the CEO of Interbank, whose sales volume by this time had slipped below Visa's, took advantage of the situation by declaring: "It appears they have begun to compete with their members."[57] He also attempted to establish a pattern of behavior by drawing a connection between this deal and Visa's recent travellers cheque program:

> Visa ventured into this territory once before, at least in the view of John J. Reynolds.... He criticized Visa's approach to travelers checks because he saw Visa, unlike the Master Charge travelers cheque program, fashioning itself into a financial services entity of its own in potential competition with members.[58]

Frederick Hammer, Senior Vice President of Chase Manhattan Bank and a Visa USA Board member, responded by urging the membership to see the Penney deal as a necessary special case, and not a change in policy or a significant threat:

> It is clear this is a good thing. It is the most efficient way to handle authorizations and value transfer. Since his costs are lower, the merchant can price his products lower. And the banks in the system get their cards used more, and that is always their goal.[59]

He admitted that in this case, Visa USA was acting more like their competitor, but he urged the members not to be concerned about this in the long run:

> I personally wouldn't mind competing with a trade association. I haven't found one yet who could do anything as well as we do. It's a lot like competing with the government.[60]

However, the controversy surrounding the deal continued to ferment "speculation and suspicion" amongst the members, and it was the "principal topic of conversation, both public and private, at the [1979] bank card convention."[61] Hock himself attempted to calm the situation

[56] Kutler, 'Visa Sued on Price Fixing, Penney Tie'.

[57] Jeffrey Kutler, 'Interbank's Reynolds Calls Visa, Penney Tie Harmful to Banking', *American Banker* (14 September 1979), p. 15.

[58] 'Usurpation Feared'. Although I do not have the space to discuss Visa's travellers cheque program here, it should be noted that Reynold's comments are strangely ironic. It was the Master Charge travellers cheque program that fashioned itself into a separate entity in competition with its members, and this was precisely the reason they were sued by Citibank. Visa's travellers cheque program worked just like the card: each bank could issue their own cheques, and the Visa mark guaranteed acceptance at any Visa member bank. Although the BofA saw the Visa travellers cheques as competition to their own, it was the issuing member banks that were competing with BofA, not Visa itself.

[59] 'Usurpation Feared'.

[60] 'Usurpation Feared'. This comment must have annoyed Hock considerably as he never wanted Visa to resemble a governmental bureaucracy.

[61] Jeffrey Kutler, 'Visa, Penney Tie Seen Unique, Not Threat', *American Banker* (17 September 1979), p. 3.

by assuring the members that the contract was unique, and that Visa did not "intend to compete with banks for merchant business." Unfortunately, he also could not resist challenging their assumption that banks, as they currently existed, would continue to be necessary in the world of electronic value exchange:

> Little in the history of bank cards has aroused such diversity of opinion, or opinion more passionately held and emotionally conveyed. All we did was supply Penney with the ability to interface with the BASE (authorization and clearing) systems.... Penney has no need for a merchant bank and in fact, insists that it not have one. Many bankers refuse to realize that many large retailers can do everything a bank can do, and often better. Penney is more sophisticated in data capture and data communications than any bank I know of. Penney's computer system is three or four times more powerful than the entire BASE system. They don't want bankers to come in and train their clerks. They have the crazy idea that they can train their clerks better than bankers can.[62]

This was exactly the kind of statement that worried the member banks most. Visa was *their* association, and Hock was *their* employee. Why was he questioning the need for an acquiring bank? Why was he designing a system that could eventually make them unnecessary? Why was he continually enabling firms from competing industries, such as the new Merrill Lynch Cash Management Accounts, to participate in the system?[63] Whose side was he on?

9.3 Hock's Departure

In the minds of the members, the JC Penney deal, even more than the Entrée card, made it obvious that Hock had a very different view of his role, and the role of the central organization, than they did. From their perspective, NBI had been created to solve a very specific problem, and Hock was simply their employee, a subordinate who existed to help them coordinate and make profits on their credit card programs. But these events now made it seem that Hock was following his own agenda, building his own empire that might someday threaten their own.

The worries of the larger and more powerful member banks quickly turned into open animosity towards Hock, often vented through the trade press. Hock responded in kind, but always with his customary touch of wit. One exemplary exchange occurred in a 1980 *Business Week* article:

> "Visa was created for the limited purpose of licensing a trademark and serving as a clearinghouse," says Hans H. Angermueller, senior vice-president of Citicorp.... Now, Angemueller says, Visa is becoming a "Frankenstein" that is working against

[62] Hock, quoted in Kutler, 'Visa, Penney Tie Seen Unique'.

[63] The Cash Management Account (CMA) offered consumers an investment account that functioned almost like a demand deposit account, complete with cheques and a Visa card, but with the potential for much higher rates of return than could be legally offered by a bank. For a history of the CMA, see Nocera, pp. 149–163. I will return to the CMA and its relation to the Visa network in Chapter 10.

> the best interests of the banks that gave it birth. Hock's response: "Citibank defines banking as Citibank. We don't."[64]

Hock's repartee might have quieted some of his critics, but he was still subject to his Board of Directors. They controlled his salary and his budgets, and if they felt he was following his own interests more than theirs, they could replace him. So far, he had been able to maintain control over the Board by appealing to those who shared his vision, and leveraging the credibility he earned by building BASE I and II, creating IBANCO, and implementing the name change. But in the minds of most Board members, the system was now "fixed," and even those who shared his vision were beginning to resent his increasingly abrasive style.

Hock tried to reassure the Board and the membership that he was not building an empire and had no ulterior motives. He would often say that "Visa is just a group of ordinary people extraordinarily committed to banking."[65] Visa was simply their coordinator and consultant, helping them cope with the radical changes occurring in the US banking industry, and Hock was still their humble servant. However, a number of developments during the early 1980s made it appear that Hock was indeed building a worldwide empire, and was more interested in his own agenda than the needs of the US members.

9.3.1 Signs of Empire Building

As discussed in Chapter 7, the Visa USA system experienced a period of explosive growth following duality and the name change, but by the late 1970s and early 1980s, the international system was also experiencing phenomenal expansion. By 1980, the Visa card was honored by more than 3 million merchants in 150 different countries, and several existing national systems had recently converted all their cards to Visa. The worldwide sales volume had climbed to $45.7 billion, nearly 40 percent of which came from issuers outside the US, and the 64 percent growth rate in Europe seemed paltry compared to Asia and Latin America, both of which were doubling their sales each year.[66]

Managing this rapid expansion was becoming difficult, as Visa International had only 500 employees, a mere 125 of which did something other than "clerical operations." By 1980, they had opened new offices in Miami (for Latin America), London (for Europe, the Middle East and Africa), and Singapore (for Asia/Pacific), but these were still rather small, with an average of ten employees each, so nearly all services were provided by the Visa USA staff. Although this allowed the US members to remain in control of the system's evolution, it also made the system seem more "American" than international. The leaders of Eurocheque and Eurocard, Visa's main competitors in Europe, often remarked that "Every time a Visa card is used in Europe, a cash register rings in the United States!"[67]

[64] 'The Iconoclast Who Made Visa No. 1', *Business Week* (22 December 1980), p. 44–46.
[65] Nocera, p. 314.
[66] Dougherty, pp. 1–4.
[67] Cleveland, p. 17; Dougherty, p. 2.

However, by this time, Hock was less concerned with the desires of the US members than he was with creating a new kind of decentralized "transnational" organization. In contrast to the approach taken by most multinational corporations, Hock wanted to divest power from the center, allowing each region to address their unique needs, create their own local policies, assess their own fees, and provide the specific services needed by their members. To achieve this, Hock re-organized Visa International in 1981 into five semi-autonomous regions, each of which had their own Board of Directors, bylaws, and operating regulations. The regions were then given seats on the International Board, which now handled inter-regional issues only, according to their percentage of the worldwide sales volume. Services such as financial analysis and product development were still provided by Visa USA's staff, but as these regions grew, those services would increasingly be provided by the regional offices.[68]

The re-organization required a bitter struggle between Hock and the US Board members, as it would effectively reduce their direct power over the worldwide system. The other regions were growing faster, and their sales would eventually overtake the US region, giving them more seats on the International Board. Hock saw it as a necessary move to make Visa a truly transnational system, and he was enamored with the idea of creating an organization with a stature similar to the United Nations, with himself as its leader and spokesperson. However, the US members saw it as yet another confirmation that he was pursuing his best interest and not theirs.

Signs of empire-building were also evident on the technical side, particularly in what became known as the "long-range plan." Visa's switch to ACP and IBM hardware for BASE I had given them the necessary capacity to keep up with the growing worldwide volumes, but Hock wanted to expand the telecommunications network across the globe and build the greatest transaction processing system the financial services world had ever seen. He and the technical staff developed an extensive wish list, which went so far as to include a private communications satellite, totaling $150 million USD. Since Visa had no public stock, and no disposable capital or assets to speak of, they turned to their members, asking them to buy a series of special notes to be paid back over ten years with eleven percent interest. The members actually had little choice; those who elected not to buy the notes were forced to pay a special "Capital Fee" that amounted to the same price as buying the notes.[69]

Although the members did purchase the first round of notes totaling $62 million, Visa's CFO noted that they felt the whole process was "too cavalier," and that without a detailed prospectus regarding how the money would be spent, the notes amounted to "an enormous 'Trust me' which translated into something else for each member."[70] They feared that Hock would use this to further enlarge his growing empire, and their suspicions were confirmed when Hock began acquiring several member-owned processing organizations on the verge of

[68] Dougherty, pp. 5–9.
[69] Cleveland, p. 24.
[70] Ibid., p. 30.

collapse, as well as a manufacturer of POS transaction switching equipment.

The international re-organization and the long-range plan both became warning signs that Hock was beginning to formulate an empire and assume a role that went far beyond what the members had ever envisioned. But the most controversial sign of all came from something that normally draws little attention: office space.

9.3.2 101 California: Headquarters for an Empire

In 1980, a new sleek high-rise office building was being planned at one of the most prestigious locations in San Francisco's financial district. It would be at the intersection of Market and California streets, overlooking not only the Bay, but also all the other firms populating that powerful section of the West Coast's banking capital. Visa's CFO wrote that it perfectly "dovetailed with Dee's vision of the future for Visa International," and Hock signed a lease for the entire forty-fourth floor. Unfortunately for him, he once again neglected to mention it to his Board.[71]

He spent the next three years turning the new space into a "showplace which included only 29 offices, enormous hallways and meeting rooms, and open sitting areas that communicated that Visa had 'arrived', and that it was a player in the world scheme of things."[72] Since the building was near-circular, Hock created a floor plan reflecting the four corners of the globe, with four sitting areas decorated to reflect each region's culture. He collected paintings, antiques, carpets, and even a bust of his most-adored leader, Marcus Aurelius, to adorn the offices. He designed a board room with state-of-the-art audio/visual equipment and integrated translation booths. It was a masterpiece of design, an office space fit for a major player in the worldwide banking industry, and that was exactly the problem.

In 1983, the Directors arrived for their first meeting in the new space, and they were aghast. Cleveland recalled their reactions:

> To them, the floor looked more like an opulent, antique showroom, not what they would have deemed proper for Visa. However, the boardroom and executive area at each Director's respective head office were just as plush, or in most cases more, than Dee's 101 creation. Unfortunately, the Directors did not share Dee's grand vision for Visa. Most had tired of Dee's methods for getting what he wanted from them. Each at some point had felt outsmarted, outwitted, outmaneuvered, and intellectually humiliated by Dee's unrelenting negotiating style. After all, most of them were lawyers, had gone to the finest schools, and had all manner of degrees and credentials, whereas Dee had merely attended junior college for a few quarters and a business career that in no way was comparable to theirs. In other words, they were tired of Dee's disrespect, and 101 California seemed to rub salt into some very old wounds.[73]

[71] Cleveland, pp. 33–36. See also Nocera, pp. 312–313.
[72] Cleveland, p. 34.
[73] Ibid.

Although the Directors were upset over the money spent on this new space, they were even more upset about the *message* the space communicated about who Hock thought he was. Hock was their employee, but his offices made him look like their equal. The offices also looked like the headquarters of a major multi-national holding company, not a simple switch for credit card transactions.

Furthermore, they realized that Hock had never sought their approval to build this new "pleasure palace in the sky," nor even to commit to the lease he had signed for the floor. They also wondered how much of the $62 million they recently loaned Visa for their long-range plan was used to build it. As Cleveland put it, "The proverbial straw was on the camels back, his front legs were buckled under, and the back legs were beginning to wobble."[74]

9.3.3 The King is Dead. Long Live the King

From that point on, Hock lost control over the board, and his detractors began to treat him with open disdain at Board meetings. Brian Ruder, one of Visa's technical staff, recalled one particular incident at a Board meeting he was asked to attend:

> I walked into the Board meeting and [Hock] is sitting there, and [a Director] was sitting next to him... they opened the meeting, and [the Director] opened the Wall Street Journal and lit a cigar, and I just couldn't believe it... it was just to say to Dee, "fuck you, I don't like you." Some of them had such disdain for Dee—they hated him.[75]

Even his stalwart supporters began to turn on him. At a Board meeting in early 1984, Hock was met with resistance at nearly every turn. Several times before he was able to break such resistance by threatening to resign if they did not follow his plans. He tried the tactic again, but this time it backfired:

> Bob Mitchell, the Chairman of the U.S. board and long-time Dee-supporter, paused and then responded to Dee's resignation proposal, "Thank you, Dee, for accepting the present circumstances and stepping down. We hereby accept your resignation, effective the day we appoint your successor, whom we would like your assistance in finding." Dee was momentarily in shock. The resignation tactic had failed. In fact, by all appearances, the decision to replace Dee had been a consensus decision reached well before the meetings had ever begun. An era was over.[76]

[74] Cleveland, p. 36.

[75] Ruder interview. Russell also told a story of a Director reading the paper during Hock's speech, and a similar incident appears in Nocera, p. 314. Ruder added that although the board disliked Hock's methods, they did respect the results he had achieved.

[76] Cleveland, p. 44. Cleveland was not actually in the room, so his account is based on what he heard later. Nocera's account, which comes mostly from interviews with the Directors, is very similar, but has Hock threatening to quit at a meeting in Greece over a denied raise in his salary (Nocera, p. 314). Russell agreed with Nocera's account, but noted that the location was actually Venice. Visa's official history claims that Hock decided to leave on his own accord, and that the Chairman "pleaded" with him to stay (Chutkow, p. 234). In his autobiography, Hock simply wrote that "in 1984, the curtain came down on my performance as CEO" (Hock, *One From Many*, p. 251).

A few months later, Chuck Russell was appointed as the new President and CEO of Visa International and USA, and Dee Hock left the system he had nurtured for fourteen years to tend his ranch on the California coast.[77] The long-range plan was cut back to fit within the funds already raised, and the 101 California lease was soon terminated, forcing the company to relocate their headquarters to the less-glamorous town of San Mateo.[78]

Several sources remarked that the transition eventually needed to happen. Hock had pushed the members too far, too fast, and he was unwilling to wait for them to realize he was right about such things as the asset card and electronic value exchange. Visa needed to evolve from a dynamic, entrepreneurial company to a more stable service organization that could dependably run an increasingly popular worldwide payment system. Debit cards and electronic value exchange would come in time, but for now, the roles that Hock and Visa would play in the future of the payment system had been unequivocally redefined.

9.4 Conclusion

In this chapter, we have seen how the roles played by Hock and the central Visa organization had to be negotiated and worked out through a series of power struggles. These struggles were evident in the members' reactions to the asset card, the JC Penney deal, and signs of Hock's empire building. In the end, Hock did not have enough power to sustain his desired role, and was eventually forced out of the organization in 1984.

In the next chapter, I will reflect upon the Visa case, deriving three lessons that extend the existing literature on payment systems, cooperative networks of competitors, and networks and social boundaries.

[77] Michael Weinstein, 'Visa Elects Russell as President, Chief Executive', *American Banker* (25 May 1984), p. 3.

[78] Cleveland, pp. 46–47. As noted in Chapter 7, Visa was able to keep their San Francisco mailing address by renting a post office box at the nearby San Francisco Airport.

Chapter 10

Conclusions: Three Lessons

The preceding chapters have been a rather long and detailed answer to the seemingly simple question posed in the introduction: how does my Visa card work? That is to say, how is it that I, living in California, can use a small piece of plastic, issued by a bank in New York that I have never visited, to purchase goods and services from merchants I have never met, or obtain local currency from a machine owned by a completely different bank, in 170 different countries around the world? How is it that a merchant can accept this piece of plastic containing a particular logo, and within seconds obtain a guarantee of payment, regardless of where the issuer is located, what time it is there, or what currency unit is used for the account? How is it that the different institutions representing the merchant and the cardholder are able to coordinate their actions, even when they reside in different countries? And how has this come to feel so "normal," or even "natural?"

Although it may feel somewhat natural today, we have seen that the creation and development of the sociotechnical system that makes this all possible was anything but that. Any payment system, even simpler ones based on commodity money or specie, should be viewed as a particular *achievement* requiring an *explanation*. To borrow a phrase from Collins, a mature payment system is like a ship within a bottle: it is sometimes difficult for us to imagine that it was once simply a pile of sticks.[1] This dissertation has recaptured and explained in detail how those metaphorical sticks became the system known as Visa today.

Throughout the dissertation, I have made use of the existing Science and Technology Studies literature where appropriate. In Chapter 4, I emphasized that the BankAmericard system was a *sociotechnical* one, arguing that their problems were both technical and organizational. In Chapter 5, I asserted that authorization had become a "reverse salient" that was holding back the development and growth of the overall system. In Chapter 6, I showed how Hock and the NBI staff "engineered" (in a wholly non-pejorative sense) cardholder expectations with the facsimile drafts, just as much as they engineered the computers and telecommunications of the BASE II system. In Chapter 7, I again emphasized the sociotechnical nature of the system by

[1] H M Collins, 'The Seven Sexes: A Study in the Sociology of a Phenomenon', *Sociology* 9:2 (1975).

observing that their rapid expansion in the 1970s was fueled by their organizational changes, but ultimately enabled by their technical upgrades. In Chapter 8, I showed how the debates surrounding the encoding of the card were politically-shaped, even down to the specific knowledge claims about security. In Chapter 9, I described how Hock and the members fought over his and Visa's proper role, and how Hock was ultimately pushed out of the organization as a result.

The existing literature did indeed prove helpful for analyzing the details of the case, but I now want to end by examining how the case might help extend the literature in return. In this concluding chapter, I will offer three "lessons" that can be teased out of this case. The first concerns the seemingly insignificant technology of a mark, and how it interacts with authority and value-exchange rules in the development and adoption of most payment systems. The second investigates how operating rules might successfully be formed and maintained within a cooperative network of competitors, and how these networks must develop a capacity for resolving disputes about their cooperative work, as well as the interpretation of those rules. The third delves into the ways in which transactional networks interact with existing social boundaries, and how the design of the network and gateways can either reinforce or undermine those boundaries.

10.1 Value Flows According to a Mark

During my interviews with Visa's pioneers, I was struck by how often they spoke about the importance of Visa's marks: the blue-white-and-gold bands design; and the name "Visa." They considered the marks to be not only key contributors to Visa's success, but also the most important asset Visa owned, even more important than the BASE I and II software. The regional Visa organizations spend vast amounts each year developing awareness of, and goodwill towards, those marks, which have subsequently become some of the most recognized in the world.[2] We have long been accustomed to products and services being branded, but it may seem odd at first that the way we *pay* for those has become seemingly branded as well. However, this is not a new phenomenon: marks have played an important role in payment systems throughout history, and will likely continue to do so in the future.

10.1.1 Marks

Marks, as I will define them, first appear in the shift from commodity money to coins. In a commodity money payment system, a particular raw material, such as shells, beads, pelts, or gold, is used as a primary trading currency. This commodity money (sometimes referred to as "full-bodied money"), still typically has some kind of intrinsic value for the participants, though that value may be more symbolic than practical. More importantly, the commodity

[2] Chutkow.

money is not issued, marked, or guaranteed by any kind of authority. It is simply a commodity that is plentiful enough to enable trade, but scarce enough to maintain sufficient value.[3]

For example, in a commodity money system, a buyer might offer a lump of gold in exchange for goods or services, but that lump would not bear any seal or mark that attested to its weight or fineness. It is simply a lump of metal ore, and no authority guarantees or stands behind it. Thus, the transaction implies some amount of risk for the seller: the lump may not be of sufficient weight, the gold might be impure, or it might not be gold at all, just some metal that merely looks like it.

Risk is endemic to all economic transactions, but a seller will typically seek to reduce that risk as much as possible by leveraging some source of knowledge the seller already trusts. One source might be the buyer himself; if the seller knows and trusts the buyer through some other social link, the seller may accept the buyer's claims as to the weight and fineness, especially if the social link affords the potential of retribution. However, if the buyer is unknown, the seller might consult the testimony of an authority, such as an expert skilled in weighing and assaying. Alternatively, the seller might perform those experiments herself using technologies developed by those skilled experts.

As early as the seventh century BCE, the Lydians began transforming the lumps of electrum they used as commodity money into something new: a *coin*. What made these "coins," as opposed to uniform lumps of metal, was the presence of a *mark*, the Lydian State seal, which attested to the ingot's weight and fineness.[4] Through the use of their mark, the Lydian State was, in essence, guaranteeing these properties of the coin, providing a new and presumably trusted source of knowledge for participants in the payment system. Because the mark belonged to a commonly-recognized and trusted authority, the participants could treat the coins as conceptual black boxes: things that could be simply counted instead of weighed and assayed.[5]

However, this black-boxing was a two-edged sword: it did not eliminate the risks as much as shift them to new locations. Participants may trust the authority and its mark, but the coin could be shaved after it was marked, the mark could be forged, or the authority itself may be corrupt. Indeed, authorities often systematically reduced the fineness of their coins in order to generate a *seignorage*, that is, a profit based on the difference between the metal content and face value of the coin. If those administering the payment system do not develop techniques to mitigate these risks, participants may quickly lose trust in the authority and its mark, rendering its guarantees or assurances essentially meaningless. Without trust, sellers and buyers will cease to participate in the payment system, and without continued exchanges, the payment system will quickly collapse.[6]

[3] Humphrey; Richardson; Chandler.

[4] Einzig, p. 217.

[5] Einzig notes that there is some disagreement as to whether the Lydian merchants accepted the coins by tally only, but that seems to have been the intent of the issuer.

[6] Following Latour, we might say that a payment system is essentially a performance, in which each value exchange recreates and sustains it. When trust fails and exchanges cease, the social links dissolve, and the system

10.1.2 Marks and Value Exchange Rules

Notes, like coins, also typically bear the marks of a commonly-recognized authority, which could be the State or some other reputable organization.[7] What makes a note different from an ordinary piece of paper is not the shape, color, or sign of denomination: it is the set of marks, the authority behind them, and the trust participants place in that authority.

However, in the case of notes, the authority guarantees something different than in the case of coins: the particular *rules of value exchange*. Recall that representative notes are only claims upon another substance (usually specie) the issuing authority currently uses to represent value. To redeem that other substance, the note must be presented to the issuer in a certain manner. The rules for how that is accomplished are tied up with the issuer's mark, and are often printed on the note itself. Incidentally, value-exchange rules are also included on notes in a fiat money system (where the notes themselves are the direct representation of value) through a performative declaration such as these notes are "legal tender for all debts public or private." Thus, a set of marks often represents a corresponding set of rules that govern final value exchange, and it is those rules that the authority "stands behind" or guarantees. If the participants within the payment system trust that authority, and believe those rules to be enforceable, the rules and their corresponding marks can provide a significant source of trust during a transaction.

10.1.3 Rules and Guarantees

It should now be easier to see why personal cheques, especially those issued in America, engender little trust. There are two principal reasons. First, although a cheque may bear the marks of the issuing bank, it does not have to. Anything can be a cheque, provided it contains the necessary information and the words "pay to the order of." Even if it does bear the marks of an issuer, those marks are known only within a limited geographic area. Merchants outside the area do not typically recognize the issuing bank as an authority, and thus have no way of knowing if the cheque is even a valid financial instrument. Second, even when the mark is present and recognized, it cannot act as a source of trusted knowledge, because the value-exchange rules that govern cheques are actually defined by a different authority (common law, or the Universal Commercial Code) and they guarantee nothing for the depositor. A cheque is only a *tentative* claim on funds, subject to inspection of the document and the account to which it is supposedly tied. Thus, the seller is limited to the same sources of knowledge discussed in the commodity money example: the seller's social links with the buyer; or the testimony of a trusted authority, which in this case would be a cheque authorization or guarantee system, if one is available.

evaporates. See Strum and Latour; and Latour, *Reassembling the Social*.

[7] This is most commonly a bank, but other organizations can issue more limited-use notes as well. For example, when I was working at Disneyland in 1987, the company began issuing "Disney Dollars," which are still accepted in most domains controlled by the Disney company.

Travellers cheques provide an interesting counter-example. They also bear the marks of a non-governmental issuing organization, but that organization typically goes to great lengths to establish itself as an authority amongst its network of recipients. Furthermore, the value-exchange rules represented by the issuer's marks actually *guarantee* payment, primarily because a travellers cheque is closer to a representative note than a conventional cheque. Recipients typically believe those rules to be enforceable, as they are written into formal business contracts subject to contract law. In an interesting twist, the retribution that was previously offered by the seller's social link to the buyer is now transfered to the recipient's social link with the issuer. Even if the travellers cheque was a forgery, the recipient can still seek restitution from the issuer, provided the recipient can demonstrate that the authentication rules were followed.

10.1.4 Marks and Payment Cards

As we come to payment cards, our analytical framework should now help us understand how the marks played a fundamental and powerful role in the adoption and development of these systems. Like other payment devices, the cards also bear the marks of an issuing organization, and in the Visa case, this issuer was initially the Bank of America (BofA). As discussed in Chapter 3, the BofA was widely recognized as a trusted authority within California, but in order to expand the system nationally, it needed to enlist the help of other banks that were already established as authorities in their local areas. In contrast to cheques, the BofA required that all cards use the same, consistent marks, and since the marks consumed the entire face of the card, all cards effectively looked alike. Because the local authority had introduced those marks to their merchants, the merchants would readily accept any card bearing those marks, regardless of issuer. Similarly, because the local authority had "trained" its cardholders to look for the marks on merchant signs, cardholders could easily identify participating merchants, regardless of acquirer. This enabled the BankAmericard system to expand geographically in a way that eluded the nascent Interbank system, whose common "i" mark was barely noticeable next to the issuer's more-dominant marks.

Furthermore, like travellers cheques, the value-exchange rules represented by the BankAmericard (and eventually Visa) marks guaranteed payment, provided the merchants followed their part of the rules. Electronic communications and the eventual replacement of human authorizers with programmed logic were the key to offering this guarantee of payment, which completely eliminated the merchant's risk. Merchants no longer needed to seek trust through their social links with cardholders, because the cardholder was no longer the "buyer." From the merchant's perspective, the "buyer" of a Visa transaction is actually the acquiring member bank, and ultimately the Visa system itself. Indeed, before the widespread adoption of point of sale terminals, the member banks would speak of "buying paper" from merchants, as opposed

to processing their transactions.[8] Thus, merchants can confidently accept Visa cards from customers they have never met, regardless of the customer's nationality, native language, or native currency unit.

This becomes most evident when we consider what the marks say about the cardholder. Just as the mark on the Lydian coin attested to the metal's weight and fineness, the mark on your Visa card attests to your creditworthiness.[9] Even if the issuer did not perform an adequate credit check on you, the merchant can still consider you as such, because Visa is the actual "buyer." When you are issued a Visa card, Visa "stands behind" you in a very real sense, because any transaction processed under the mark is ultimately guaranteed by the Visa system, even if you default or overdraw, and even if the issuing organization goes out of business.

Even if Visa is not the ultimate buyer, the marks on your Visa card may still attest to your general trustworthiness. For example, before the widespread use of cheque guarantee systems, American merchants commonly asked those paying with a cheque to provide a major credit card as a secondary proof of identity. What made that piece of plastic an acceptable form of economic identification was not the embossed account number or expiration date: it was the marks, the authority behind them, and the trust the merchants placed in that authority.

Lastly, trust in the authority behind a payment card system, its corresponding marks, and the value-exchange rules represented by those marks, can falter just as easily as it can with other kinds of payment devices. Network and processing outages, billing errors, security breaches, financial organization failures, or excessive (or well-publicized) fraud can quickly erode participants' trust in the system. As we saw in Chapter 5, the BASE I authorization system was designed from the beginning to deliver a reliable service, and the capacity enhancements discussed in Chapter 7 enabled BASE I to remain dependable throughout Visa's period of explosive growth. These are only a small part of Visa's overall focus on dependability, which I have discussed in detail elsewhere.[10]

10.1.5 Application

In summary, *value always flows according to a mark*.[11] The mark (or set of marks) is established by an authority, and only those who recognize and trust that authority will continue to participate in the payment system. Through the mark, the authority attests to or guarantees something, which can be properties of the instrument, or the rules by which the instrument can be exchanged for value. If those rules guarantee payment, and if participants believe them to

[8] For example, see Kenneth V Larkin, 'Dealer Paper in the Age of the Consumer', *Banking* 63:4 (October 1970), p. 32.

[9] Or, your ability to pay in general. The mark on your debit card might not be related to a credit line, but from the merchant's perspective, it attests to a guarantee of payment.

[10] David L Stearns, 'In Plastic We Trust: Dependability and the Visa Payment System', ⟨URL: `http://www.sociology.ed.ac.uk/finance/Papers/StearnsDIRC06.pdf`⟩.

[11] This phrase comes from the insights of Tom Honey, who took the time to explain the importance of Visa's marks during our interviews.

be enforceable, they can provide a significant source of trust that can offset many of the risks merchants face in accepting payments, or at least shift them to other parties, where they can be addressed in a consolidated manner.

There is one important implication of this lesson that should be briefly mentioned here: payment systems are not neutral conduits through which value flows. Rather, value flows according to a mark, which represents certain rules that govern not only how the value is exchanged, but also who pays for the service, who pays for fraud, how disputes are resolved, and who gets to alter and interpret the rules.[12] In most cases, those rules will favor one group over another, and occasionally may be designed to restrict access. Although we rarely think about these things when paying for goods and services, payment systems have real and often profound effects upon the ways in which value is exchanged.

10.2 Rules, Disputes, and Interpretation

Our second lesson builds upon the first: if value flows according to rules established by an authority, and represented by a mark, then how are those rules created and maintained when that "authority" is actually a cooperative network of competitors? How do those competitors balance out their own individual interests to ensure that the rules do not unfairly favor one group of members over another? Even more importantly, what happens when those competitors begin to have disputes over transactions, especially those that call into question what those rules actually mean? How do they resolve those disputes in order to sustain the network?

10.2.1 Maintaining the Rules

In the Visa system, these rules are encoded into organizational bylaws and operating regulations at each level of the system. For example, Visa USA has its own bylaws and operating regulations, but they are subject to those established by Visa International. Similarly, other national organizations such as Carte Bleue maintain their own bylaws and operating regulations, but they are subject to those of Visa Europe, which are in turn subject to those of Visa International. In keeping with Dee Hock's organizational philosophy, his basic principles, which are encoded in the bylaws, are dictated from the international level, but the day-to-day operational rules, which are encoded in the operating regulations, are pushed as far down the hierarchy as possible to allow for local variation and adaptation.[13] Thus, the international operating regulations dictate only those aspects that are necessary to enable interregional interchange, and ensure consistent use of the worldwide marks.

[12] This is similar to the point made in Howells and Hine. Their focus was on the process of designing a new electronic funds transfer system, but this corollary suggests that all past or existing payment systems also have inherent politics. Winner has also discussed the possibility of inherently political technologies in Langdon Winner, 'Do Artifacts Have Politics?', *Daedalus* 109 (1980).

[13] See section 4.3 beginning on page 68.

When NBI formed in 1970, Hock established the first set of operating regulations and required members to abide by them "as they now exist or are hereafter modified."[14] These rules served as a coordination mechanism that enabled the competitors to cooperate just enough to provide a ubiquitous service none could have supplied alone. As discussed in Chapter 4, further changes to those were then discussed and approved by the advisor groups, which were made up of representatives from various member banks. The advisor groups proved to be particularly useful for two reasons. First, they allowed NBI's staff to "take the temperature" of the membership on potentially controversial issues without causing widespread or intractable debate. Second, if the advisor groups liked the ideas, they would then act as advocates for them, making it seem that the ideas came from the general membership and not NBI's staff. Regulation changes were then voted upon by the Board of Directors (or one of their subcommittees), which was carefully structured by Hock to provide adequate representation for all types of member banks within the system (those within a given geographic area, those smaller than a certain size, etc.).

When IBANCO formed in 1974, the international operating regulations were copied from those used by NBI, but any rules that were specific to a given national system were removed. Further changes were then discussed and voted upon by the International Board, which was structured to grant blocking power to each of the regions, ensuring that no one region or bank could dominate policy decisions. When the system was reorganized in 1981 into five semi-autonomous regions, each regional organization developed their own set of operating regulations that were based upon the International set, but were further enhanced to deal with region-specific services and issues. Each region was allowed to develop their own mechanisms for altering those rules, and most emulated the arrangement of advisor groups and a representative Board of Directors structured to ensure blocking power.

Thus, Visa's primary method for balancing out the individual interests of the competitors was a kind of democratic representation that strove to provide various member groups with sufficient *blocking power* at the Board level.[15] In theory, member groups were forced to form coalitions and make compromises in order to make significant alterations to the rules. In practice, most of the important changes made during the first decade of the system were actually driven by Hock, who often used the large number of Directors and advisors to his own advantage.[16] However, the representative structure helped the members *feel* like their concerns were being heard, their needs were being met, and the rules were not egregiously tipped against them.

[14] Hock, *One From Many*, p. 124.

[15] Although the advisor groups had no real policy-making power, they did exert a significant influence. If they strongly opposed a proposal, it was not likely to be put forward for a Board vote.

[16] See the quote from Russell on page 73.

10.2.2 Dispute Resolution

The operating regulations stipulated not only how interchange transactions were to be processed, but also how disputes over those transactions were to be resolved within the system. Dispute resolution is a critical feature of any cooperative network of competitors; without it, the competitive members would have no way to resolve disagreements that invariably arise, especially when money is involved. From the beginning, Visa defined a chargeback and arbitration process that would provide a method for resolving these disputes in as fair and equitable a manner as possible.[17]

The process has remained essentially the same throughout the years, though it has been modified slightly based on experience. The process typically begins when a cardholder complains about a particular transaction. This complaint can be raised for a number of reasons: the cardholder did not conduct the transaction (or does not wish to admit to it[18]); the transaction was keyed incorrectly or was processed multiple times; the merchandise was "not as advertised"; the merchandise was never received; was defective or damaged in shipment; etc.[19] If the issuer thinks the complaint is valid, the issuer then submits a chargeback transaction through BASE II. Because the original transaction had already been cleared and settled, the issuer essentially puts through a compensating transaction, "charging back" the acquirer for the original amount plus a penalty fee. The acquirer then has a certain number of days to research the transaction on their side, after which they can either close the case or "represent" the transaction through BASE II. Originally, the operating regulations allowed the transaction to bounce back and forth a few times, but today, the issuer and acquirer must resolve the dispute after one cycle, or submit it to Visa for arbitration.

Initially, the group that handled chargeback disputes within Visa was the same group responsible for maintaining the operating regulations. In the 1970s, interchange volumes were still rather low, and disputes were less common, so the arbitration process was simply a side job of those maintaining the regulations. As interchange volumes increased, so did the disputes, and Visa reacted by developing a more complex, formalized arbitration process, and staffing positions within the operating regulations area to manage it.[20]

This group's primary responsibility is to shepherd cases through Visa's formal arbitration process. When two members cannot resolve a chargeback dispute amongst themselves, they submit it to this group, along with all the supporting documentation, and a fee to discourage

[17] Information about dispute resolution comes from interviews with Tindal, Baum, and Kollmann.

[18] Several interviewees indicated that male cardholders would occasionally conduct morally questionable transactions that they later regretted, and instead of admitting to them, would claim their card had been lost or stolen.

[19] Consumer protection laws in the United States allow cardholders to refuse payment for goods that are "not as advertised" or defective if they are purchased within a certain, limited geographic range from the cardholder's residence. In other countries this may not be a legitimate complaint.

[20] Unfortunately, statistics on the number of disputes raised for arbitration are not available to the public, but interview sources indicated that they were relatively few, perhaps a dozen per week, less than a tenth of a percent of all chargebacks. The dispute resolution staff actually spends more time educating and fielding general questions from the membership.

frivolous cases. In the early years, the cases were reviewed and decided by an ad-hoc committee composed of mid-to-upper-level managers, but today the dispute resolution group itself acts as the "court." In most cases, the applicable rules are fairly clear and unambiguous, but the cardholder/issuer and merchant/acquirer testimonies conflict. The court then does their best to review all the available evidence and decide which party is correct.

However, in some cases the dispute is based on differences of *interpretation*; that is, the evidence might be clear, but the two parties have different opinions as to whether their actions adhered to the rules. For example, the rules require merchants to compare the signature on the sales draft with the signature on the card, but what constitutes a match? Two signatures that might seem similar-enough to a merchant might seem completely different to an issuer or cardholder. Initially, the rules did not clarify what it meant for two signatures to "match," and the members of the arbitration committee were forced to make their own judgments on a case-by-case basis. As the number and dollar value of these kinds of disputes increased, this quickly became untenable, and the rules were eventually modified to require only that the *content* of the signatures matched, and not the hand in which they were written.[21]

Although they may be rare, these differences in interpretation are always a possibility in rule-based systems. As Wittgenstein argued, a rule in itself, cannot fully specify what it means to follow or not follow the rule.[22] The interpretation of a rule is necessarily bound up with participants' attempts to follow it, disputes over those attempts, and judgments made by a commonly-recognized authority as to whether those attempts were correct or not. In this case, Visa acts as the commonly-recognized authority, establishing a canonical interpretation of the rules through their arbitration decisions.

However, as discussed in the signature example, Visa also has the luxury of modifying rules that are seen as too vague or problematic. Visa's ultimate goal is to minimize disputes and arbitration cases, so if particular rules cause too many interpretation disputes, Visa suggests changes to clarify them. Although the operations committee must approve those changes, they have little incentive to resist, as disputes reduce the overall efficiency of the system.

Additionally, the dispute resolution group takes an *active* role in educating the membership about the chargeback and arbitration rules. The operating regulations are actually quite detailed, and currently fill multiple printed volumes, so it is not uncommon for a member to be unfamiliar with certain sections. The resolution group helps to educate the members on the entire process, including their rights and obligations. They also send out regular communications, informing the members of new rules and their implications, or recent arbitration decisions that might help clarify how the rules should be interpreted.

[21] Honey interview. Note that this rule typically applies to fraudulent transactions involving a lost or stolen card that has since been recovered. If the contents of the signatures on the sales drafts do not match the card, the merchant or acquirer must absorb the chargebacks. If the contents do match, even if they are in a completely different hand, the issuer must absorb the loss.

[22] Ludwig Wittgenstein, *Philosophical Investigations*, Translated by G E M Anscombe, 3rd edition (Oxford: Blackwell, 1958), § 201–202.

The decisions handed down by Visa's "court" are typically final, but under certain circumstances, such as high-value transactions, the loser is allowed to appeal the decision. There are currently several layers through which the case can be appealed, with the Board of Directors acting as the final, supreme arbiter. Interestingly, Visa's arbitration process is not entirely binding; the members still retain the right to bring suit against one other in their country's legal system if they do not like the result of the arbitration.

It should be noted here that although Visa's chargeback and arbitration process is well designed, it is by no means perfect. Powerful members can and do abuse the mechanism for their own gain. Smaller banks have been known to pay questionable chargebacks from intimidating issuers without a fight, primarily because they lacked the resources to investigate the matter and collect the necessary documentation. Those reviewing the cases are humans, and as such, can never be completely impartial. Lastly, members have occasionally felt that the arbitration process did not produce the correct result, and have chosen to continue fighting in the civil court system.

10.2.3 Application

Nevertheless, it is clear that the dispute resolution process was a key sustaining mechanism for the Visa system, and might be an important feature of any stable cooperative network of competitors. Whenever competitors engage in cooperative work, disputes between the participants will inevitably arise, especially if there is money or power to be gained or lost. Because these networks tend to be rather fragile, a network that wishes to survive must develop a capacity for resolving these disputes.

Through examining the Visa case, we can see that certain features of their capacity proved useful, and thus might also appear in other stable networks. First, the members surrendered some autonomy by agreeing to follow a set of rules, jointly defined and maintained, as they existed and were thereafter modified. Second, those same rules also dictated how disputes were to be resolved, allowing Visa to modify the process as they gained experience. Third, members were given a relatively simple mechanism by which they could resolve disputes amongst themselves, if possible. Fourth, when that proved impossible, they could then use a formal arbitration process, administered by a commonly-recognized and impartial authority. Fifth, that authority also sought to minimize future disputes by clarifying vague rules and actively educating the members about the process and current interpretations. Last, the members had an appeal process for highly-significant disputes that would ultimately lead to a jury of their peers, the Board of Directors.

10.3 Networks and Boundaries

So far, we have discussed how value flows according to rules represented by a mark, and how Visa created, maintained, and resolved disputes about those rules within a cooperative network of competitors. Now we need to shift our focus towards the technological networks that actually make that value "flow," and how they interact with established social boundaries.

10.3.1 Making Value Flow

As opposed to other kinds of digitized content, such as audio files, value does not flow across a communications network like water flows through a pipe. There are no electronic pennies that stream across the wire from one node to another. Instead, value is exchanged by offsetting the balances of accounts that are linked together into a hierarchical network in response to standardized instructional messages. Let us build this up, step by step.

A single bank is perhaps the simplest example of a network of accounts. All the accounts maintained by a given bank are in effect linked together. When one account holder wishes to transfer money to another, the payer does so by sending some sort of standardized message indicating at least the source, target, and amount. The funds are then transfered by simply debiting the payer's account and crediting the payee's.

However, the banks themselves are also linked together through complex networks of accounts. Smaller banks often maintain accounts with a correspondent bank in a larger money center. Some of those larger banks also maintain correspondent accounts with banks in other countries to enable foreign exchange. Banks belonging to a cooperative clearinghouse typically maintain accounts with a common clearing bank. In the US, members of the Federal Reserve System are also required to maintain accounts with their respective district bank. This complex network of accounts allows the banks to move value quickly and safely over long distances by simply sending standardized messages that instruct the recipient to debit one account and credit another. Value still "flows," but only in a metaphorical sense.

A cheque or a payment card sales draft is actually a similar kind of standardized message: it instructs the issuer to debit the payer's account and credit the bearer's account upon presentment. If the bearer is a different bank, the paying bank merely sends another standardized message to the clearing bank to debit its account and credit the other bank's account.

Interestingly, all of this existed long before the banks adopted digital computers. The standardized messages were encoded onto paper, which naturally limited the speed at which they could travel, but once they were received, value moved through simple bookkeeping entries. Thus, the "revolution" in payment systems that occurred in the 1960s and 1970s was not solely enabled by the digitization of *accounts*, for those were already stored in an easily-manipulable form and never left the domain of the bank. Rather, it was primarily enabled by the digitization of those *standardized messages*.

When Visa truncated the paper sales drafts and began clearing transactions electronically

through BASE II, those messages could now travel anywhere in the world almost instantly. This enabled Visa to clear and settle transactions for the entire system overnight, greatly accelerating the flow of value, as well as reducing the handling and float costs incurred by the slow movement and difficult reconciliation of paper drafts.

However, it also introduced a new possibility. From a technical perspective, any organization that maintained accounts could become a node on this network, regardless of geographic location or industry. In addition to banks, this could include investment brokers, national retail chains, universities, telecommunications providers, transit systems, utility companies, and a host of others. In other words, many different kinds of organizations could now suddenly act like a "bank," accepting deposits and providing electronic payment services.

This was a frightening potential for bankers, who had grown accustomed to the protections afforded by the strict regulatory boundaries that separated them from other industries. The question then became, will the Visa network be extended in a way that respects these existing boundaries, or transgresses them? Would these other types of organizations be allowed to participate in the payment system, and if so, on what terms?

10.3.2 Boundaries

I should pause at this point to define what I mean by a "boundary." Boundaries, as I am using them here, are conceptual demarcations erected by actors to define a given social world. They might define the extent of a workgroup, a discipline, an organization, an industry, or a political state. Boundaries establish who is inside and who is outside, and as such, designate who is allowed to participate in certain activities, who is allowed to access or contribute resources, and who is allowed to influence policy decisions. They can be crossed, but not without effort or cost. However, it is important to realize that these boundaries are neither natural nor static: they are actively erected and defended, and often challenged or transgressed. New boundaries may appear, existing ones may shift, or established ones may disappear altogether.[23]

The main lesson of this section is that transactional networks are one mechanism by which these boundary changes may occur. Transactional networks link various actors together to perform work, often on behalf of an external user, and when those networks cross established social boundaries, they can begin to undermine them in subtle ways, making them more porous, and eventually rendering them more cosmetic than restrictive. Often, this causes actors to question those boundaries and ultimately remove them. Of course, this does not always occur. Powerful actors may defend existing boundaries and limit or control the ways in which networks are allowed to cross over them, especially when those boundaries provide a significant political or economic advantage.

[23]For the concept of boundary-working, see Thomas F Gieryn, 'Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists', *American Sociological Review* 48:6 (December 1983); and MacKenzie, *Inventing Accuracy*, pp. 413–417.

10.3.3 Physical and Logical Structure

When discussing networks and boundaries, it is also important to make the distinction between a network's *physical* and *logical* structure. The physical structure refers to the arrangement of nodes and how they are connected to one another. Any given node has one or more direct connections to other nodes, but those nodes may also connect to still other nodes, creating indirect paths between nodes that are not directly connected.

A *gateway* is a special kind of node that allows a network's physical structure to respect established social boundaries. A gateway stands between two networks, each of which can be controlled by different parties. The gateway is responsible for bridging the two networks, forwarding messages between them and translating the protocols if necessary.[24] The key point is that in order to get to the other network, one must go through the gateway, and submit to its terms.

In contrast to the physical structure, the logical structure of a network refers to the way in which *the user conceives it*. The degree to which the physical and logical structures differ is largely determined by how visible or noticeable the gateways are. For example, on the Internet, a message may route through several nodes and across several gateways before reaching its destination, but the connections are entirely seamless. From the user's perspective, the source and destination nodes appear to be connected *directly*. In contrast, when one uses an ATM in the US that is owned by a bank other than the one who holds the source account, the gateways are made quite visible through the warning and levy of additional "foreign" transaction fees.

Thus, when a transactional network, such as Visa, connects organizations across existing social boundaries, the designers have a choice as to whether the physical structure will respect the boundaries, *and* whether to make that physical structure evident or seamless to the user. Even if established boundaries are respected using gateways, the users will consider those boundaries cosmetic, or not even notice them at all, if the gateways are not made obvious. Seamless connections will eventually undermine these boundaries, and if they are not actively defended or reconstructed, they may be called into question and ultimately removed.

10.3.4 Transgressing and Defending Boundaries

The Visa case provides us with a few examples of how this might work out in practice. As we have seen, enabling connections to any organization that held value was at the heart of Hock's vision of electronic value exchange. He wanted to allow cardholders to access not only a line of credit issued by a bank, but also any pool of value they might possess, including demand deposits and investments. The question was not *if* the network should be extended to include these, but *how*.

[24] In computer networking, the node is called a "gateway" if it translates protocols, or a "router" if it merely forwards the messages without conversion. Since I am dealing more with questions of boundaries than communication, I will simplify the discussion by using the term gateway only.

Boundaries Within the Banks: Debit Cards

As we saw in Chapter 9, the boundary between deposit and credit accounts within most banks was actively defended when NBI introduced the asset card. NBI was essentially proposing that banks extend their internal BankAmericard network to include the demand deposit accounts in addition to the credit accounts. This challenged the boundary that the deposit bankers had erected to keep this strange unsecured consumer credit device away from the "real banking." In banks where this boundary was strong, the card was not adopted, and thus the networks remained unlinked.

Eventually, this boundary was breached at most banks, but not until several generations of bank management had retired. By that time, many other changes, both within the banking industry and the Visa system, had made the boundary less important, and the deposit bankers of the 1990s were less interested in defending it. Once the boundary was crossed, it began to be questioned, and in many banks today it has been completely removed. At most banks, customers can link their debit and credit accounts together, and use the latter to cover transactions when there are insufficient funds in the former.

Boundaries Between Types of Banks: the Thrifts

Although NBI was initially unsuccessful at crossing boundaries within the banks, they were successful at transgressing some of the categorical boundaries within the banking industry. When NBI formed in 1970, the initial bylaws restricted membership to commercial banks only, primarily because all the existing BankAmericard licensees were commercial banks. Shortly thereafter, Hock began pushing to open the membership to mutual savings banks, savings and loan associations, and credit unions as well. These other types of banks, collectively referred to as "the thrifts," were more restricted in their activities than commercial banks, and some types were organized as not for profit. They primarily accepted savings deposits and made relatively safe loans, such as mortgages. As such, they were not subject to the same regulations as the commercial banks, and most importantly, could offer higher rates of interest on deposits.

However, the commercial banks could offer their customers a key service that the thrifts could not: third-party payment instruments.[25] One could deposit money into a thrift, but getting it back out again required a trip to the branch to obtain cash. It was far more convenient for customers to put their money into a commercial bank account, where they could write cheques against it, or use a credit line attached to their BankAmericard.

In the early 1970s, many thrifts began to eliminate this advantage by offering what became known as "Negotiable Order of Withdrawal" (NOW) accounts. These were interest-bearing savings accounts that acted like demand deposit accounts, complete with a negotiable payment

[25]Third-party payment instruments can be transfered from the account holder to a party other than the source bank, such as a merchant. See James F Nielsen, *Banking Terminology*, 3rd edition (Washington DC: American Bankers Association, 1989).

instrument that looked and acted just like a traditional cheque. If the thrifts were allowed to join NBI, they could then issue BankAmericards as well. This would undermine the boundaries between commercial and thrift banks, especially in the depositor's mind, making them simply cosmetic.

Hock eventually got his way, and the thrifts were allowed to join NBI provided their state charters allowed it.[26] The physical structure of the transactional network was then extended over the existing boundaries separating the various types of banks, destroying them completely. There was still a legal distinction between the various types, but from the depositor's point of view, the boundary was no longer restrictive. Thrift accounts could now participate in the cheque and BankAmericard payment systems just as well as commercial bank accounts could.

Boundaries Between Financial Organizations: CMAs

A similar story can be told about the boundaries between different kinds of financial organizations, particularly those between banks and investment brokerages. Like the thrifts, investment accounts were not subject to the same regulations as bank accounts, and thus could offer a return that was closer to the current market interest rate. However, investment accounts naturally incurred a higher risk, and like the thrifts, they typically did not offer any kind of third-party payment instruments.[27]

In the early 1970s, several investment organizations began offering what became known as "money market funds," which pooled together customer deposits and invested them into relatively safe, short-term securities, such as Treasury bills.[28] However, these funds differed from ordinary investments in that customers could withdraw their capital incrementally, on demand. When Fidelity introduced their version in 1974, they completed the service by issuing cheques, just as the thrifts had done with NOW accounts.

In 1977, Merrill Lynch took this one step further when they unveiled their new "Cash Management Account" (CMA).[29] This rather creative service combined a traditional investment portfolio with a money market account. Investment dividends and sale proceeds were automatically "swept" into the money market account where they began earning market-rate interest immediately. But the most interesting aspect of the CMA was that it came not only with cheques, but also *a Visa card*. Visa transactions debited the cash in the money market fund, or an automatically funded margin account secured by the investments if not enough cash was available. It was exactly what Hock had been advocating with his asset card concept, except it

[26] 'NBI Amends By-Laws to Permit Membership by Savings Banks', *American Banker* (18 October 1973), p. 1. Russell noted that this was partially due to fears of antitrust litigation. If they had denied membership to the thrifts, they could have been sued for being exclusive.

[27] An *investment account* is one used primarily to purchase or sell securities such as stocks and bonds, though it may also contain excess, uninvested cash. Unlike a traditional bank savings account, cash in a US investment account is not protected by the Federal Deposit Insurance Corporation (FDIC), and is thus considered more risky.

[28] Nocera, pp. 75–88.

[29] Ibid., pp. 155–164.

was being provided by an investment broker and not a commercial bank.

Although the member banks agreed to let the thrifts join, they were not about to let investment brokers become members. In 1977, the US economy was experiencing a high rate of inflation, and observant consumers noticed that the interest rates on their commercial savings accounts were actually *less* than the rate of inflation. The buying power of their money was evaporating, but the commercial banks were prohibited by law from offering a higher rate, or any interest at all on demand deposits. The CMA offered consumers something that looked and acted just like a commercial bank account, but provided a rate of return that would not only preserve their money's buying power, but also increase it.

Since Merrill could not join Visa as a member, it contracted with City National Bank and Trust (CNBT) of Columbus, Ohio, to act as their gateway into the network. CNBT built software to bridge the two networks, forwarding both BASE I authorization messages as well as cleared transactions from BASE II. From the perspective of the Visa system, CNBT was the issuer and thus responsible for settling the transactions, but once the Visa settlement was complete, CNBT's computer systems immediately settled again with Merrill.

The member banks won the battle to make the physical structure respect the existing boundaries, but they eventually lost the war because the logical structure did not. CNBT made the gateway completely seamless, and as such, made it appear to cardholders that Merrill was directly connected to Visa's transactional network. A Merrill CMA customer had no idea that CNBT was involved, nor did they need to. From the depositor's point of view, the boundaries between commercial banks and CMA-style accounts became so porous that they began to question why the boundaries existed at all.

Today, the regulatory boundaries between commercial banks and investment brokers no longer exist, but Visa USA's bylaws still prohibit the brokers from becoming members. To gain access to Visa, most brokers have since purchased some type of bank (commonly an Industrial Loan Company), which can become a Visa member, and provide a seamless gateway between Visa and their own account networks. The physical structure of the Visa network still respects the boundaries that have been defended by the member bankers, but the logical structure has rendered them mostly meaningless for the consumer.

Boundaries Between Industries: JC Penney

For a final example, let us consider the JC Penney deal that was already discussed at length in Chapter 9. In order to get JC Penney to accept Visa cards, which would eventually force the other large national retailers to accept them as well, Visa had to allow JC Penney to connect directly into the transactional network, bypassing an acquiring bank. This was not simply a direct technological connection for submitting transactions: JC Penney was setup with a bank identification number and settlement account, just like any other member bank. Thus the network was physically extended over the existing boundary that separated banks from large

retailers.

As we saw, the Board reluctantly agreed to honor this deal, but JC Penney remained the only US retailer ever to participate directly in the settlement network. After signing the deal, the Board amended the bylaws to prohibit this kind of incursion from ever happening again. Thus, they effectively defended the boundary between banks and retail organizations, which still exists to this day. All other retailers must go through a member bank, which acts as a gateway to the settlement networks. As opposed to the gateways used by the investment brokers, these are not seamless. The discount fee keeps the gateway noticeable, reminding the merchants that the boundary is restrictive and actively defended.

10.3.5 Application

In summary, the networks that make value move in the Visa system were sometimes designed to respect existing social boundaries, and other times to transgress them. In the case of the asset card, the deposit bankers defended their boundaries for a time, keeping the network away from their accounts. In the case of the thrifts, the physical structure of the network punctured the boundaries between the different kinds of banks, eventually rendering them cosmetic. In the case of the CMA, the physical structure respected the boundaries, but because the gateway was so seamless, the logical structure eventually undermined them. In the case of JC Penney, the bankers successfully defended the boundary, forcing other large retail merchants to go through an acquirer, which is essentially a noticeable gateway.

This interaction between network structure and boundaries should apply to the study of other of cases as well. Whenever a transactional network is created to link disparate actors together across existing social boundaries, the designers will necessarily have to decide if the physical structure will respect those boundaries or not. If gateways are used, they can be configured to be seamless, or noticeable and costly to traverse, and that will have profound effects on how the users of the network perceive those boundaries. Powerful actors may seek to defend their boundaries by limiting traffic through the gateway, or using it to extract fees. On the other hand, seamless connections will tend to undermine those boundaries, calling them into question, and jeopardizing their future existence.

10.4 Epilogue

In this final chapter, we have seen how the Visa case can help us extend the literature on payment systems, cooperative networks of competitors, and the relationship between transactional networks and social boundaries. I discussed how value always flows according to a mark, and how those marks can mitigate risk when they guarantee something, such as value-exchange rules that are backed by a commonly-recognized and trusted authority. I described Visa's mechanisms for establishing, maintaining, and adjudicating the operating rules that govern

how their network functions, suggesting that other such stable cooperative networks of competitors will likely develop similar mechanisms. Lastly, I noted how transactional networks can be designed to either respect or transgress existing social boundaries, and the ways in which seamless gateways can eventually undermine and call those boundaries into question.

These general lessons build upon the rich and detailed history that consumes the bulk of this dissertation, and it is about this history that I wish to make my concluding remarks. Part of what attracted me to this case in the first place was the way in which Visa has become such an integral yet unexamined part of our daily lives. Visa may be one of the best known brands in the world, but few have any idea what Visa is, how it is structured, or how it functions. Most of us think of our Visa cards the same way we think of a water faucet, a light switch, or a telephone. We no longer notice the systems lying behind these access devices because they have faded into the background, blended in with the task itself, and effectively become invisible.

When we hand our Visa card to a merchant, we rarely think about the magnetic stripes, integrated circuits, terminals, telecommunications, computers, and software programs that make electronic value exchange possible. Nor do we think about the people who determined how to make those cards machine-readable, how to build inexpensive point of sale terminals, how to connect merchants and banks throughout the world, how to authorize thousands of transactions each second, and clear and settle hundreds of millions of dollars worth of purchases each night in multiple currencies. Nor do we think about those who worked tirelessly to standardize magstripe and transaction formats so that devices and computer systems could exchange information easily and efficiently. Nor do we think about the organizational structures that allow the 20,000 Visa members around the world to cooperate just enough to coordinate their work, resolve their disputes, and determine future policy. We rarely think about this because, at least most of the time, it simply "works."

Forty years ago, Californians would crowd around the checkout to see someone pay with a BankAmericard. Today, it has become so commonplace that we are shocked and somewhat annoyed when the person in front of us at the grocery store begins the agonizingly slow process of writing a paper cheque. The first time I put my Visa card into an ATM in a foreign country and received local currency, it seemed like magic. Now I rarely think twice before boarding an international flight with nothing but a Visa card for payment. We are no longer amazed when a merchant far from home swipes our card through a terminal and receives an authorization, a *guarantee of payment*, a few seconds later. We have just come to expect it, as we have with countless other technological innovations.

However, this is precisely the reason we historians of technology do what we do. We tell the stories of how technological systems came to be so that we can not only appreciate them more, but also recapture how they have altered the ways in which we think and act. In essence, our job is to make the invisible technologies visible again, if only for a moment.

Bibliography

Abbate, Janet, *Inventing the Internet* (Cambridge, MA: MIT Press, 1999).

Alvesson, Mats and Per Olof Berg, *Corporate Culture and Organizational Symbolism: An Overview* (Berlin: Walter de Gruyter, 1992).

Anderson, Stuart et al., 'Down in the (Data)base(ment): Supporting Configuration in Organisational Information Systems' (April 2003), Paper presented at the EPSRC DIRC Conference.

Arthur, Brian W, 'Competing Technologies and Economic Prediction', in: MacKenzie, Donald and Judy Wajcman, editors, *The Social Shaping of Technology*, 2nd edition (Maidenhead: Open University Press, 1999), pp. 106–112.

Baker, Donald, 'Shared ATM Networks: The Antitrust Dimension', *Federal Reserve Bank of St. Louis Review* (November/December 1995), pp. 5–17.

'Bank Cards Push for the Big Stores', *Business Week* (27 September 1976), pp. 107–108.

'Bank Cards Take Over the Country', *Business Week* (4 August 1975), pp. 44–54.

'Bank Credit-Card and Check-Credit Plans', Federal reserve system report (Federal Reserve System, July 1968).

Barnes, Barry, David Bloor and John Henry, *Scientific Knowledge: A Sociological Analysis* (Chicago: University of Chicago Press, 1996).

Baxter, William F, 'Bank Interchange of Transactional Paper: Legal and Economic Perspectives', *Journal of Law and Economics* 26:3 (October 1983), pp. 541–588.

'Beating the New Credit Cards', *Business Week* (11 August 1973), pp. 120–121.

Becker, Howard S, 'Culture: A Sociological View', *Yale Review* 71 (1982), pp. 513–527.

Bernard, Jules, 'New Directions in Bankcard Competition', *Catholic University Law Review* 30 (1980), pp. 65–102.

Berry, Henry V, 'The Credit Card and the Debit Card', in: *Critical Issues in Bank Cards* (Washington DC: American Bankers Association, 1980), pp. 102–123.

Bertalanffy, Ludwig von, *General System Theory: Foundations, Development, Applications* (New York: George Braziller, 1968).

Besnard, Denis and Gordon Baxter, 'Human Compensations for Undependable Systems', Technical report CS-TR-819 (School of Computing Science, University of Newcastle upon Tyne, November 2003).

Bezroukov, Nikolai, 'Open Source Software Development as a Special Type of Academic Research (Critique of Vulgar Raymondism)', *First Monday* 4:10 (October 1999).

'The Big-Name Credit Cards and How They Compare', *Changing Times* 27 (September 1973), pp. 25–28.

Bloch, Marc, 'The Watermill and Feudal Authority', in: MacKenzie, Donald and Judy Wajcman, editors, *The Social Shaping of Technology*, 2nd edition (Maidenhead: Open University Press, 1999), pp. 152–155.

Bodnar, John, 'Power and Memory in Oral History: Workers and Managers at Studebaker', *Journal of American History* 75:4 (March 1989), pp. 1201–1221.

Borland, Katherine, "'That's not what I said": Interpretive Conflict in Oral Narrative Research', in: Perks, Robert and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998), pp. 320–332.

Bradford, Terri and Stuart E Weiner, 'Who's Processing Your Payments?', Payments system research briefing (Federal Reserve Bank of Kansas City, August 2005), ⟨URL: `http://www.kc.frb.org/FRFS/PSR/PSR-BriefingAug05.pdf`⟩ – visited on 6 October 2005.

Brandel, Roland E and Joseph E Terraciano, 'Legal and Regulatory Environment', in: *Critical Issues in Bank Cards* (Washington DC: American Bankers Association, 1980), pp. 124–152.

Brooke, Phillip, 'Quigley Describes Operation of Omniswitch; Urges Cooperation in National Development', *American Banker* (11 August 1971), pp. 6–7.

Brooke, Phillip, 'BankAmericard Schedules April Start for 24-Hour National Card Authorization', *American Banker* (21 August 1972), p. 1.

Brooke, Phillip, 'IBANCO, Ltd., Organized to Manage Worldwide BankAmericard Program', *American Banker* (20 September 1974), p. 1.

Brooks, Frederick P, *The Mythical Man Month: Essays on Software Engineering*, Anniversary edition (1975; Reprint, Reading, MA: Addison-Wesley, 1995).

Buchan, James, *Frozen Desire: The Meaning of Money* (New York: Farrar, Straus and Giroux, 1997).

Buchanan, R A, 'Theory and Narrative in the History of Technology', *Technology and Culture* 32:2 (April 1991), pp. 365–376.

Burns, Arthur F, *The Ongoing Revolution in American Banking* (Washington DC: American Enterprise Institute for Public Policy Research, 1988).

Burns, Tom and G M Stalker, *The Management of Innovation* (London: Tavistock, 1961).

Callon, Michel, 'The Sociology of an Actor-Network: The Case of the Electric Vehicle', in: Callon, Michel, John Law and Arie Rip, editors, *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World* (Basingstoke: MacMillan, 1986), pp. 19–34.

Callon, Michel, 'Society in the Making: The Study of Technology as a Tool for Sociological Analysis', in: Bijker, Weibe E, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987), pp. 83–107.

Campbell-Kelly, Martin, *From Airline Reservations to Sonic the Hedgehog : A History of the Software Industry* (Cambridge, MA: MIT Press, 2003).

Campbell-Kelly, Martin and William Aspray, *Computer: A History of the Information Machine* (New York: Basic Books, 1996).

'Can the Stores Trump Bank Credit Cards?', *Business Week* (30 January 1971), pp. 62–64.

'Card Plans Show Big Gains', *Burroughs Clearing House* 52:5 (February 1968), pp. 15–16.

Carr, E H, *What is History?* (New York: Penguin Books, 1961).

Carrington, Mark, Philip W Langguth and Thomas D Steiner, *The Banking Revolution: Salvation or Slaughter? How Technology is Creating Winner and Losers* (London: Pitman Publishing, 1997).

Ceruzzi, Paul E, *A History of Modern Computing* (Cambridge, MA: MIT Press, 1998).

Chandler, Alfred D, *The Visible Hand: The Managerial Revolution in American Business* (Cambridge, MA: Harvard University Press, 1997).

Chandler, Lester V, *The Economics of Money and Banking*, Ninth edition (1948; Reprint, New York: Harper and Row, 1986).

'The Charge-It Plan That Really Took Off', *Business Week* (27 February 1965), pp. 58–63.

Cheney, Julia and Sherrie Rhine, 'How Effective Were the Financial Safety Nets in the Aftermath of Katrina?', Discussion paper (Payment Cards Center of the Federal Reserve Bank of Philadelphia, January 2006), ⟨URL: `http://www.philadelphiafed.org/pcc/HurricaneKatrinaJan06.pdf`⟩ – visited on 2 Feb 2006.

Chown, John, *A History of Money: From AD 800* (London: Routledge, 1994).

Chutkow, Paul, *Visa: The Power of an Idea* (Chicago: Harcourt, 2001).

Clark, Burton R, 'The Organizational Saga in Higher Education', *Administrative Science Quarterly* 17 (1972), pp. 178–184.

Cleveland, Tom, 'The Visa History: Tom Cleveland's Perspective' (April 1999), Unpublished compilation of personal stories by Tom Cleveland, former CFO of Visa International.

Coghlan, Richard, *The Theory of Money and Finance* (London: Macmillan, 1980).

Collins, H M, 'The Seven Sexes: A Study in the Sociology of a Phenomenon', *Sociology* 9:2 (1975), pp. 205–224.

Collins, H M, 'In Praise of Futile Gestures: How Scientific is the Sociology of Scientific Knowledge?', *Social Studies of Science* 26 (1996), pp. 229–244.

Collins, Harry, 'The TEA Set: Tacit Knowledge and Scientific Networks', *Social Studies of Science* 4:2 (April 1974), pp. 165–186.

Colton, Kent W and Kenneth L Kraemer, editors, *Computers and Banking: Electronic Funds Transfer Systems and Public Policy* (New York: Plenum Press, 1980).

Competition and Innovation in the Credit Card Industry at the Consumer and Network Level, Hearings before the Subcommittee on Financial Institutions of the Committee on Banking, Housing, and Urban Affairs, United States Senate, 106th Congress (25 May 2000).

The Consumer Credit Protection Act Ammendments of 1977, Hearings on HR 8753 before the Subcommittee on Consumer Affairs of the Committee on Banking, Finance, and Urban Affairs, United States House of Representatives, 95th Congress, First Session. (March 1977).

Copeland, Duncan G, Richard O Mason and James L McKenney, 'Sabre: The Development of Information-Based Competence and Execution of Information-Based Competition', *IEEE Annal of the History of Computing* 17:3 (1995), pp. 30–57.

'Credit Cards: A Study Of Consumer Attitudes and Usage' (1975), Survey conducted for National BankAmericard Incorporated by Market Facts Incorporated.

Critical Issues in Bank Cards (Washington DC: American Bankers Association, 1980).

'A Data Stripe Puts Speed in Credit Cards', *Business Week* (16 September 1972), pp. 68–69.

David, Paul A, 'Path Dependence and the Quest for Historical Economics: One More Chorus of the Ballad of QWERTY', Discussion Papers in Economic and Social History 20 (University of Oxford, November 1997).

Davies, Glyn, *A History of Money: From Ancient Times to the Present Day* (Cardiff: University of Wales Press, 1994).

Desmonde, William H, *Magic, Myth, and Money: The Origin of Money in Religious Ritual* (New York: Free Press of Glencoe, 1962).

'Digital and Banking' (1976), Promotional brochure produced by Digital Equipment Corporation.

'Don't Just Pay that Charge Account Bill. Read It!', *Changing Times* (February 1973), pp. 37–40.

Dosi, G, 'Technological Paradigms and Technological Trajectories: A Suggested Interpretation of the Determinants and Directions of Technical Change', *Research Policy* 11:3 (June 1982), pp. 147–162.

Dougherty, J Stewart, 'Visa International: The Management of Change', Harvard Business School Case Study (1981), Prepared under the supervision of Robert G Eccles, Jr.

Douglas, Mary, *Risk Acceptability According to the Social Sciences* (London: Routledge and Kegan Paul, 1985).

Drury, Tony and Charles W Ferrier, *Credit Cards* (London: Butterworths, 1984).

Dryer, Davis, 'The Limits of Technology Transfer: Civil Systems at TRW, 1965–1975', in: Hughes, Agatha C and Thomas P Hughes, editors, *Systems, Experts, and Computers: The Systems Approach in Management and Engineering, World War II and After* (Cambridge, MA, 2000), pp. 359–384.

Dyer, Davis, *TRW: Pioneering Technology and Innovation Since 1900* (Boston: Harvard Business School Press, 1998).

Ederer, Rupert J, *The Evolution of Money* (Washington DC: Public Affairs Press, 1964).

Einzig, Paul, *Primitive Money in its Ethnological, Historical and Economic Aspects*, 2nd edition (1949; Reprint, Oxford: Pergamon, 1966).

Essinger, James, *Electronic Payment Systems: Winning New Customers* (London: Chapman and Hall, 1992).

Essinger, James, *The Virtual Banking Revolution: The Customer, the Bank and the Future* (London: International Thomson Business Press, 1999).

Evans, David and Richard Schmalensee, *Paying with Plastic: The Digital Revolution in Buying and Borrowing*, 2nd edition (Cambridge, MA: MIT Press, 2005).

Evans, David S, 'Bank Interchange Fees Balance Dual Demand', *American Banker* 166:18 (26 January 2001), p. 17.

Evans, David S, 'Will Retailers Stampede to Drop Signature Debit?', *American Banker* 169:1 (2 January 2004), p. 10.

'Facts About BankAmericard', National BankAmericard Incorporated Booklet (October 1975).

Fazar, Willard, 'The Origin of PERT', *The Controller* (December 1962), pp. 598–602, 618–621.

Felgran, Steven D, 'Shared ATM Networks: Market Structure and Public Policy', *New England Economic Review* (January/February 1984), pp. 23–38.

Felgran, Steven D, 'From ATM to POS Networks: Branching, Access, and Pricing', *New England Economic Review* (May/June 1985), pp. 44–61.

Fenstermaker, J Van and Donald Perry, 'An Examination of a Charge Card System', *Journal of Bank Research* 2:1 (1971), pp. 9–13.

Fernelius, Leonard W and David Fettig, 'The Dichotomy Becomes Reality: Ten Years of the Federal Reserve as Regulator and Competitor' ⟨URL: `http://minneapolisfed.org/pubs/ar/ar1991.cfm`⟩ – visited on 26 August 2006.

Fisher, John et al., *Bank Cards* (Washington DC: American Bankers Association, 1980).

Francis, G Jay and Susan M Siegel, *Principles of Banking: An Overview of the Financial Services Industry*, Seventh edition (Washington DC: American Bankers Association, 2001).

Frazer, Patrick, *Plastic and Electronic Money: New Payment Systems and Their Implications* (Cambridge: Woodhead-Faulkner, 1985).

Friedlander, Peter, 'Theory, Method and Oral History', in: Perks, Robert and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998), pp. 311–319.

Friedman, Andrew L, *Computer Systems Development: History Organization and Implementation* (New York: John Wiley and Sons, 1989).

Fry, Maxwell J et al., *Payment Systems in Global Perspective* (London: Routledge, 1999).

Galanoy, Terry, *Charge It: Inside the Credit Card Conspiracy* (New York: GP Putnam's Sons, 1980).

Galbraith, John Kenneth, *Money: Whence it Came, Where it Went*, Revised edition (1975; Reprint, Boston: Houghton Mifflin, 1995).

Gerdes, Geoffrey R and Jack K Walton, 'The Use of Checks and Other Noncash Payment Instruments in the United States', *Federal Reserve Bulletin* (August 2002), ⟨URL: `www.federalreserve.gov/pubs/bulletin/2002/0802_2nd.pdf`⟩.

Gerson, Elihu M, 'Reach, Bracket, and the Limits of Rationalized Coordination: Some Challenges for CSCW', Chapter for the forthcoming *Resources, Co-Evolution, and Artifacts: Theory in CSCW*, edited by Ackerman et al..

Gieryn, Thomas F, 'Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists', *American Sociological Review* 48:6 (December 1983), pp. 781–795.

Goldberg, Lawrence G, 'The Effect of State Banking Regulations on Bank Credit Card Use', *Journal of Money, Credit and Banking* 7:1 (February 1975), pp. 105–112.

Good, Barbara Ann, *The Changing Face of Money: Will Electronic Money Be Adopted in the United States?* (New York: Garland, 2000).

'Great Wolf Lodge Rides the RFID Wristband Wave' ⟨URL: `http://www.pdcorp.com/crowd-control/case-study-great-wolf.html`⟩ – visited on 13 September 2006.

Grele, Ronald J, 'Movement Without Aim: Methodological and Theoretical Problems in Oral History', in: Perks, Robert and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998), pp. 38–52.

Gunn, Janet Varner, *Autobiography: Towards a Poetics of Experience* (Philadelphia: University of Pennsylvania Press, 1982).

Hagiu, Andrei, 'Cash, Credit, or Car?' ⟨URL: `http://www.marketplatforms.com/payingwithplastic/mpdperspectives/`⟩ – visited on 7 September 2006.

Hardstone, Gillian, Luciana D'Adderio and Robin Williams, 'Standardization, Trust and Dependability' (2004), Unpublished chapter.

Hashagen, Ulf, Reinhard Keil-Slawik and Arthur Norberg, editors, *History of Computing: Software Issues* (Berlin: Springer, 2002).

Hayashi, Fumiko, Richard Sullivan and Stuart E Weiner, *A Guide to the ATM and Debit Card Industry* (Kansas City, MO: Payment Systems Research Department, Federal Reserve Bank of Kansas City, 2003).

Head, Robert V, 'Getting Sabre Off the Ground', *IEEE Annal of the History of Computing* 24:4 (2002), pp. 32–39.

Hock, Dee, 'Bank Card Industry Must Press for Legislation on National Level', *American Banker* (3 October 1973), p. 4.

Hock, Dee, 'Electronic Funds Transfer or Electronic Value Exchange?' (October 1974), Paper presented at the Federal Reserve Bank of Boston Conference, Melvin Village, New Hampshire.

Hock, Dee, 'Competition for Financial Services in an Electronic Banking Environment' (March 1975), Speech given to the National Marketing Conference of the American Bankers Association, Chicago.

Hock, Dee, 'Bank Cards: Today's Popular Song or An Unfinished Symphony?' (14 September 1976), Speech given at the American Bankers Association Bank Card Conference, San Francisco.

Hock, Dee, *Birth of the Chaordic Age* (San Francisco: Berrett-Koehler, 1999).

Hock, Dee, *One From Many: VISA and the Rise of Chaordic Organization* (San Francisco: Berrett-Koehler, 2005), second edition of *Birth of the Chaordic Age*.

Hopton, David, *Payment Systems: A Case for Consensus* (Bank for International Settlements, 1983).

Hosemann, Michael J, 'The Rationale for Electronic Banking', in: *Critical Issues in Bank Cards* (Washington DC: American Bankers Association, 1980), pp. 60–101.

'How Banking Tames Its Paper Tiger', *Business Review of the Federal Reserve Bank of Philadelphia* (May 1960), pp. 2–11.

Howells, John and James Hine, editors, *Innovative Banking: Competition and the Management of a New Networks Technology* (London: Routledge, 1993).

Hughes, Thomas P, *Networks of Power: Electrification in Western Society, 1880-1930* (Baltimore, MD: Johns Hopkins University Press, 1983).

Hughes, Thomas P, 'The Seamless Web: Technology, Science, Etcetera, Etcetera', *Social Studies of Science* 16:2 (May 1986), pp. 281–292.

Hughes, Thomas P, 'The Evolution of Large Technological Systems', in: Bijker, Wiebe E, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987), pp. 51–82.

Humphrey, David B, *Payment Systems: Principles, Practice and Improvements* (Washington DC: The World Bank, 1995).

'The Iconoclast Who Made Visa No. 1', *Business Week* (22 December 1980), pp. 44–46.

'Introduction to Entrée' (1975), Brochure on the new Entrée card produced by National BankAmericard Incorporated for its members.

Jevons, William Stanley, *Money and the Mechanism of Exchange* (London: Twentieth Century, 1875).

Johnson, Robert W and A Charlene Sullivan, 'Statistical Analysis of Bank Card Usage', in: *Critical Issues in Bank Cards* (Washington DC: American Bankers Association, 1980), pp. 1–59.

Jutilla, Donald, *Fundamentals of Bank Credit Cards* (Tacoma, WA: American Institute of Banking, Tacoma Chapter, 1973).

Katz, Michael L and Carl Shapiro, 'Network Externalities, Competition, and Compatibility', *American Economic Review* 75:3 (June 1985), pp. 424–440.

Kent, Raymond P, *Money and Banking*, 4th edition (1947; Reprint, New York: Holt, Rinehart and Winston, 1961).

Kirkman, Patrick, *Electronic Funds Transfer Systems: The Revolution in Cashless Banking and Payment Methods* (Oxford: Basil Blackwell, 1987).

Klebaner, Benjamin J, *Commercial Banking in the United States: A History* (Hinsdale, IL: Dryden Press, 1974).

Kling, Rob, 'Value Conflicts and Social Choice in Electronic Funds Transfer System Developments', *Communications of the ACM* 21:8 (August 1978), pp. 642–657.

Knight, John R, 'A Case Study: Airlines Reservations Systems', *Proceedings of the IEEE* 60:11 (November 1972), pp. 1423–1431.

Kraft, Philip, *Programmers and Managers: The Routinization of Computer Programming in the United States* (New York: Springer-Verlag, 1977).

Kuhn, Sarah, 'The Limits to Industrialization: Computer Software Development in a Large Commercial Bank', in: Wood, Stephen, editor, *The Transformation of Work? Skill Flexibility and the Labour Process* (London: Unwin Hyman, 1989).

Kuhn, Thomas S, *The Structure of Scientific Revolutions*, 3rd edition (1962; Reprint, Chicago: University of Chicago Press, 1996).

Kunda, Gideon, *Engineering Culture: Control and Commitment in a High-Tech Corporation* (Philadelphia: Temple University Press, 1992).

Kutler, Jeffrey, 'Metal Plates to Duality: The Shaping of an Industry', *American Banker* (9 September 1994), pp. 30–36, Special insert.

Laprie, JC, editor, *Dependability: Basic Concepts and Terminology in English, French, German, Italian and Japanese* (Vienna: Springer, 1992).

Larkin, Kenneth V, 'Dealer Paper in the Age of the Consumer', *Banking* 63:4 (October 1970), pp. 32–33+.

Latour, Bruno, *Science in Action: How to Follow Scientists and Engineers Through Society* (Cambridge, MA: Harvard University Press, 1987).

Latour, Bruno, 'Ethnography of a High-Tech Case: About Aramis', in: Lemonnier, Pierre, editor, *Technological Choices: Transformation in Material Cultures Since the Neolithic* (London: Routledge, 1993), pp. 372–397.

Latour, Bruno, *Aramis, or, The Love of Technology* (Cambridge, MA: Harvard University Press, 1996).

Latour, Bruno, 'Give Me a Laboratory and I Will Raise the World', in: Biagioli, Mario, editor, *The Science Studies Reader* (New York: Routledge, 1999), pp. 258–275.

Latour, Bruno, 'On Recalling ANT', in: Law, John and John Hassard, editors, *Actor Network Theory and After* (Oxford: Blackwell, 1999), pp. 15–25.

Latour, Bruno, *Reassembling the Social : An Introduction to Actor-Network-Theory* (Oxford: Oxford University Press, 2005).

Law, John, 'Technology and Heterogeneous Engineering: The Case of Portuguese Expansion', in: Bijker, Wiebe E, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987), pp. 111–134.

Lejeune, Philippe, *On Autobiography*, Translated by Katherine Leary, edited by Paul John Eakin (Minneapolis, MN: University of Minnesota Press, 1989).

Leveson, Nancy G, *Safeware: System Safety and Computers* (Reading, MA: Addison-Wesley, 1995).

Levinson, Marc, *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger* (Princeton, NJ: Princeton University Press, 2006).

Liebowitz, S J and Stephen E Margolis, 'Path Dependence, Lock-in, and History', *Journal of Law, Economics, and Organization* 11:1 (April 1995), pp. 205–226.

Losowsky, Andrew, 'I've Got You Under My Skin', *The Guardian* (10 June 2004), ⟨URL: `http://technology.guardian.co.uk/online/story/0,3605,1234827,00.html`⟩ – visited on 13 September 2006.

Lummis, Trevor, 'Structure and Validity in Oral Evidence', in: Perks, Robert and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998), pp. 273–283.

MacKenzie, Donald, 'Missile Accuracy: A Case Study in the Social Process of Technological Change', in: Bijker, Weibe E, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987), pp. 195–222.

MacKenzie, Donald, *Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance* (Cambridge, MA: MIT Press, 1990).

MacKenzie, Donald, 'How Do We Know the Properties of Artefacts? Applying the Sociology of Knowledge to Technology', in: Fox, Robert, editor, *Technological Change: Methods and Themes in the History of Technology* (Amsterdam: Harwood, 1996), pp. 247–263.

MacKenzie, Donald, *Mechanizing Proof: Computing Risk and Trust* (Cambridge, MA: MIT Press, 2001).

MacKenzie, Donald, 'A View from the Sonnenbichl: On the Historical Sociology of Software and System Dependability', in: Hashagen, Ulf, Reinhard Keil-Slawik and Arthur Norberg, editors, *History of Computing: Software Issues* (Berlin: Springer, 2002), pp. 97–122.

MacKenzie, Donald and Judy Wajcman, editors, *The Social Shaping of Technology*, 2nd edition (1985; Reprint, Maidenhead: Open University Press, 1999).

Magnis, Nicholas E, 'Descriptive Billing—A Solution to the Paperwork Problem', *ABA Banking Journal* 62:10 (April 1970).

Mandell, Lewis, *Credit Card Use in the United States* (Ann Arbor, MI: University of Michigan Institute for Social Research, 1972).

Mandell, Lewis, *The Credit Card Industry: A History* (Boston: Twayne, 1990).

Mandell, Lewis and Neil B Murphy, *Bank Cards* (Washington DC: American Institute of Banking and American Bankers Association, 1976).

Manning, Robert D, *Credit Card Nation: The Consequences of America's Addiction to Credit* (New York: Basic Books, 2000).

Marti, John and Anthony Zeilinger, *Micros and Money: New Technology in Banking and Shopping* (London: Policy Studies Institute, 1982).

Martin, Brian, 'The Critique of Science Becomes Academic', *Science Technology and Human Values* 18:2 (1994), pp. 247–259.

Martin, Brian, 'Sticking a Needle into Science: The Case of Polio Vaccines and the Origin of AIDS', *Social Studies of Science* 26 (1996), pp. 245–276.

Mason, James Eliot, *The Transformation of Commercial Banking in the United States, 1956–1991* (New York: Garland, 1997).

Mathews, H Lee and John W Slucum, Jr, 'Correlatives of Commercial Bank Credit Card Use', *Journal of Bank Research* 2:4 (1972), pp. 21–27.

Mayer, Martin, *The Bankers: The Next Generation* (New York: Truman Talley, 1997).

McAndrews, James J, 'The Evolution of Shared ATM Networks', *Federal Reserve Bank of Philadelphia Business Review* (May/June 1991), pp. 3–16.

McLeod, Robert W, *Bank Credit Cards for EFTS: A Cost-Benefit Analysis* (Ann Arbor, MI: University Microfilms International, 1979).

Meyer, Alan D and William H Starbuck, 'Interactions Between Ideologies and Politics in Strategy Formation', in: Roberts, Karlene H, editor, *New Challenges to Understanding Organizations* (New York: Macmillan, 1993), pp. 99–116.

Millard, Stephen, 'The Foundations of Money, Payments and Central Banking: A Review Essay' ⟨URL: `http://d.repec.org/n?u=RePEc:mmf:mmfc06:106&r=his`⟩.

'Money Goes Electronic in the 1970s', *Business Week* (13 January 1968), pp. 54–76.

'National Authorization Joint Feasibility Study Final Report' (29 January 1971), Provided by member of committee.

Newlyn, W T, *Theory of Money*, 2nd edition (1962; Reprint, Oxford: Clarendon, 1971).

Nielsen, James F, *Banking Terminology*, 3rd edition (Washington DC: American Bankers Association, 1989).

Nilson, Spencer, *The Nilson Report* (Oxnard, CA: HSN Consultants, 1970 to present), A twice-monthly newsletter on the payment card industry.

Nocera, Joseph, *A Piece of the Action: How the Middle Class Joined the Money Class* (New York: Simon and Schuster, 1994).

O'Brien, James A, *The Impact of Computers on Banking* (Boston: Bankers, 1968).

'Omniswtich: A Cornerstone for National Authorization Through Interbank' (14 December 1971), Description of the Omniswitch system prepared for a meeting in Atlanta, GA. Provided by Tom Schramm, VP of operations for Omniswitch.

O'Neil, Milton, 'Charge-Card Networks Are Working', *ABA Banking Journal* 66:3 (September 1973), pp. 116ff.

O'Neil, Paul, 'A Little Gift From Your Friendly Banker', *Life* 68 (27 March 1970), pp. 48–50.

Orr, Julian, *Talking About Machines: An Ethnography of a Modern Job* (Ithaca, NY: Cornell University Press, 1996).

Parker, Martin, *Organizational Culture and Identity: Unity and Division at Work* (London: SAGE, 2000).

Perks, Robert and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998).

Perrow, Charles, *Normal Accidents: Living With High-Risk Technologies* (New York: Basic Books, 1984).

Peters, Thomas J and Robert H Waterman, *In Search of Excellence: Lessons from America's best-run companies* (New York: Harper and Row, 1982).

Pettigrew, Andrew M, 'On Studying Organizational Cultures', *Administrative Science Quarterly* 24 (1979), pp. 570–581.

Pinch, Trevor J and Wiebe E Bijker, 'The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other', *Social Studies of Science* 14:3 (August 1984), pp. 399–441.

Pinch, Trevor J and Wiebe E Bijker, 'The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other', in: Bijker, Wiebe E, Thomas P Hughes and Trevor Pinch, editors, *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (Cambridge, MA: MIT Press, 1987), pp. 17–50.

Poitevent, J L, 'OCR for Credit Card Processing', *Datamation* 15:7 (July 1969), pp. 49–52.

Pollock, Neil and James Cornford, 'ERP Systems and the University as an 'Unique' Organisation', *Information Technology and People* 17:1 (2004), pp. 31–52.

Portelli, Alessandro, 'What Makes Oral History Different?', in: Perks, Robert and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998), pp. 63–74.

'The Public Appraises Bank Cards: A National Survey of Public Attitudes Toward and Use of Bank Cards' (1971), Survey conducted for National BankAmericard Incorporated by Field Research Corporation.

Raymond, Eric S, *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*, Revised edition (Sebastopol: O'Reilly, 2001).

Reason, James, *Human Error* (Cambridge: Cambridge University Press, 1990).

Reason, James, 'Human Error: Models and Management', *British Medical Journal* 320 (March 2000), pp. 768–770.

'The Recession Catches Up with Credit Cards', *Business Week* (17 February 1975), p. 47.

'The Republican's Bold Credit Card Campaign', *Business Week* (4 December 1978), pp. 33–34.

Richardson, Dennis W, *Electric Money: Evolution of an Electronic Funds-Transfer System* (Cambridge, MA: MIT Press, 1970).

Roberts, Karlene H, 'Some Characteristics of One Type of High Reliability Organization', *Organization Science* 1:2 (1990), pp. 160–176.

Roberts, Karlene H, editor, *New Challenges to Understanding Organizations* (New York: Macmillan, 1993).

Rochlin, Gene I, 'Defining 'High Reliability' Organizations in Practice: A Taxonomic Prologue', in: Roberts, Karlene H, editor, *New Challenges to Understanding Organizations* (New York: Macmillan, 1993), pp. 11–32.

Ross, Irwin, 'The Credit Card's Painful Coming-of-Age', *Fortune* 84:4 (October 1971), pp. 108–111, 150, 155–156.

Rowlingson, Karen and Elaine Kempson, *Paying with Plastic: A Study of Credit Card Debt* (London: Policy Studies Institute, 1994).

Russell, Thomas, *The Economics of Bank Credit Cards* (New York: Praeger, 1975).

Sapolsky, Harvey M, *The Polaris System Development: Bureaucratic and Programmatic Success in Government* (Cambridge, MA: Harvard University Press, 1972).

Schantz, Herbert F, *The History of OCR: Optical Character Recognition* (Manchester Center, VT: Recognition Technologies Users Association, 1982).

Schreft, Stacey L, 'How and Why Do Consumers Choose Their Payment Methods?' (October 2005), Prepared for the Federal Reserve Bank of Bostons Consumer Behavior and Payment Choice conference, October 2005.

Schulman, Paul, 'The Analysis of High Reliability Organizations: A Comparative Framework', in: Roberts, Karlene H, editor, *New Challenges to Understanding Organizations* (New York: Macmillan, 1993), pp. 33–53.

Scrutchin, Jr, Thomas W, 'TPF: Performance, Capacity, Availability', in: *IEEE Compcon 1987 Digest of Papers* (Washington, DC: IEEE Computer Society, 1987), pp. 158–160.

Simmel, Georg, *The Philosophy of Money*, Third Enlarged edition (1978; Reprint, London: Routledge, 2004).

Simmons, Matty, *The Credit Card Catastrophe: The 20th Century Phenomenon That Changed the World* (New York: Barricade Books, 1995).

Simon, Herbert A, *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*, 3rd edition (New York: The Free Press, 1976).

Siu, Lucia, 'Octopus and Mondex: The Social Shaping of Money, Technology and Consensus' (April 2002), ⟨URL: `http://www.sociology.ed.ac.uk/finance/Papers/Siu_Octopus.pdf`⟩ – visited on 4 September 2006, Social Studies of Finance working paper.

Siwiec, J E, 'A High-Performance DB/DC System', *IBM Systems Journal* 16:2 (1977), pp. 169–195.

Smith, Sidonie and Julia Watson, editors, *Getting a Life: Everyday Uses of Autobiography* (Minneapolis, MN: University of Minnesota Press, 1996).

Smith, Sidonie and Julia Watson, *Reading Autobiography: A Guide for Interpreting Life Narratives* (Minneapolis, MN: University of Minnesota Press, 2001).

Spengemann, William C, *The Forms of Autobiography: Episodes in the History of a Literary Genre* (New Haven, CT: Yale University Press, 1980).

Stallwitz, John F, 'A Market Research Study of the Attitudes and Reactions of Member and Non-Member Merchants to Bank Credit Card Plans', Masters thesis, University of California Berkeley (December 1968).

Stanley, Liz, *The Auto/Biographical I: The Theory and Practice of Feminist Auto/Biography* (Manchester: Manchester University Press, 1992).

Stanley, Liz, 'On Auto/Biography in Sociology', *Sociology* 27:1 (February 1993), pp. 41–52.

Starr, Chauncey, 'Social Benefit versus Technological Risk', *Science* 165:3899 (September 1969), pp. 1232–1238.

Staudenmaier, John M, *Technology's Storytellers: Reweaving the Human Fabric* (Cambridge, MA: MIT Press, 1985).

Stearns, David L, 'In Plastic We Trust: Dependability and the Visa Payment System', ⟨URL: `http://www.sociology.ed.ac.uk/finance/Papers/StearnsDIRC06.pdf`⟩, Paper presented at the DIRCshop Conference, 10–11 April 2006, Newcastle, UK.

Stigler, George J, 'The Division of Labor is Limited by the Extent of the Market', *Journal of Political Economy* 59:3 (June 1951), pp. 185–193.

Streeter, Bill, 'Let's Stop Using EFT Jargon, and Start Offering New Services', *ABA Banking Journal* 71:7 (July 1979), pp. 69–75, Interview with Dee Hock.

Struble, Frederick M, 'Bank Credit Cards and Check Credit Plans in the Nation and the District', *Monthly Review of the Tenth Federal Reserve District* (July–August 1969), pp. 3–9.

Strum, Shirley and Bruno Latour, 'Redefining the Social Link: From Baboon to Humans', *Social Science Information* 26:4 (1987), pp. 783–802.

Swindells, Julia, editor, *The Uses of Autobiography* (London: Taylor and Francis, 1995).

Thompson, Paul, *The Voices of the Past: Oral History*, 3rd edition (Oxford: Oxford University Press, 2000).

Thomson, Alistair, 'Anzac Memories: Putting Popular Memory Theory Into Practice in Australia', in: Perks, Robert and Alistair Thomson, editors, *The Oral History Reader* (London: Routledge, 1998), pp. 300–310.

Unsolicited Bank Credit Cards, Hearings before the Committee on Banking and Currency, United States House of Representatives, 19th Congress, First Session. (8 and 9 November 1967).

'Using Credit Cards to Buy the Groceries', *Business Week* (9 November 1974), p. 50.

Vansina, Jan, *Oral Tradition: A Study in Historical Methodology*, Translated by H M Wright (London: Routledge, 1965).

Vaughan, Diane, *The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA* (Chicago: University of Chicago Press, 1996).

Visa, 'VisaNet Fact Sheet' ⟨URL: `http://corporate.visa.com/md/fs/corporate/visanet.jsp`⟩ – visited on 26 June 2007.

Visa, 'Visa International Annual Report' (2006), ⟨URL: `http://corporate.visa.com/av/reports/corp_report.jsp`⟩.

'A Visa-Card Offensive Angers the Opposition', *Business Week* (5 September 1977), p. 31.

'Visa Debit Card Service: A Digest of Key Research Findings' (March 1977), Produced by National BankAmericard Incorporated.

'Visa Debit Card Services' (1977), Brochure on the Visa Debit card produced by Visa USA.

'Visa Dial Terminal Pilot Project Final Report' (April 1982), Produced by Visa USA.

'Visa's Climb to No. 1 in Cards', *Business Week* (16 July 1978), pp. 26–27.

Weick, Karl E, *The Social Psychology of Organizing*, 2nd edition (New York: Random House, 1969).

Weick, Karl E, 'Organizational Culture as a Source of High Reliability', *California Management Review* 29:2 (Winter 1987), pp. 112–127.

Weick, Karl E, *Sensemaking in Organizations* (Thousand Oaks, CA: SAGE, 1995).

Weinberg, Allen, 'Payments 2004: The Merchant Perspective: Exploring the Payments Acceptance Landscape', Technical report (Glenbrook Partners, August 2004), ⟨URL: `http://www.glenbrook.com/opinions/merchant-perspective-2004.html`⟩ – visited on 6 October 2005.

Weinberg, Gerald M, *An Introduction to General Systems Thinking* (New York: John Wiley and Sons, 1975).

Weiner, Stuart E, 'Electronic Payments in the U.S. Economy: An Overview', *Federal Reserve Bank of Kansas City Economic Review* 84:4 (1999).

'Why Big Stores Are Taking Outside Credit Cards', *Business Week* (3 September 1979), pp. 132–134.

Wiegold, C Frederic, 'NBI Appeals to Court for Chance to Prove Dual Membership Ban is Not Group Boycott', *American Banker* (2 January 1973), p. 1.

Winner, Langdon, 'Do Artifacts Have Politics?', *Daedalus* 109 (1980), pp. 121–136.

Witt, Ulrich, 'Firms as Realizations of Entrepreneurial Visions', Papers on Economics and Evolution 0510 (Max Planck Institute of Economics, Jena, 2005), ⟨URL: `https://papers.econ.mpg.de/evo/discussionpapers/2005-10.pdf`⟩ – visited on 2 Feb 2006.

Wittgenstein, Ludwig, *Philosophical Investigations*, Translated by G E M Anscombe, 3rd edition (Oxford: Blackwell, 1958).

Wonglimpiyarat, Jarunee, *Strategies of Competition in the Bank Card Business: Innovation Management in a Complex Economic Environment* (Brighton: Sussex Academic Press, 2004).

Zelizer, Viviana A, *The Social Meaning of Money: Pin Money, Paychecks, Poor Relief and Other Currencies* (New York: Basic Books, 1994).

List of Interviewees

All interviews were conducted by the author. Dates and locations of the interviews follow each name. All subjects were previous employees of Visa, except where noted. Job titles and departmental affiliations meant very little during this period, as the company was small, project-oriented, and subject to constant reorganization. In the early days, Hock allowed his employees to create their own job titles, and although they were eventually standardized, they followed the conventions of banking, where nearly all key players are given a variant of the title "Vice President."

Terry Abrams. 29 March 2006, telephone.

Anonymous. 21 April 2006, telephone.

Elaine Baum. 8 June 2007, telephone.

Carol Coye Benson. 13 December 2005, telephone.

Greg Bjorndahl. 22 November 2005, telephone.

Brad Boston. 24 January 2006 and 6 March 2006, telephone.

Victor Chinn. IBM Systems Engineer assigned to Visa 1985–1989. 20 September 2005, telephone.

Walter Conway. 17 May 2006, San Francisco, CA.

Tom Cleveland. 8 November 2005 and 7 December 2005, telephone.

Irwin (Win) Derman. 28 September 2005, 21 October 2005, and 11 January 2006, Milbrae, CA.

Denny Dumler. Although he eventually became a Visa employee, Dumler was previously with Colorado National Bank and the PLUS ATM association. 9 December 2005 and 6 January 2006, telephone.

Linda Elliott. 3 November 2004, Denver, CO.

Frank Fojtik. 23 May 2006, Dublin, CA.

Jeff von Gillern. 23 February 2006, telephone.

George Glasser. McKinsey consultant hired for the BASE I project. Spring 2005, telephone.

Dave Goldsmith. 28 March 2006, telephone.

Scott Harrison. 16 May 2006, telephone.

H Robert (Bob) Heller. 10 October 2005, Tiburon, CA.

Tom Honey. 26 June 2006 and 18 July 2006, San Francisco, CA.

Mick Hosken. 6 April 2006, telephone.

Perry Hudson. Chase Manhattan Bank and the American Bankers Association's card standardization committee. 12 March 2007, telephone.

Denver Huff. Involved with the BankAmericard program at First National Bank of Oregon (formerly First National Bank of Portland). 8 May 2006, Grants Pass, OR.

Mike Jones. 4 April 2006, telephone.

Don Jutilla. Managed the BankAmericard program at Puget Sound National Bank in Tacoma, Washington. 10 March 2006, telephone.

Bennett Katz. 26 October 2005, San Francisco, CA.

Ingrid Kollmann. 18 May 2006, telephone.

Scott Loftesness. 13 October 2005, Palo Alto, CA.

Lewis Mandell. Historian of Finance and Professor at the SUNY, Buffalo. 1 March 2006, telephone.

Richard Martin. 19 April 2006, Orinda, CA.

David Nordemann. 5 June 2006, San Mateo, CA.

Ron Olive. 26 January 2006, telephone.

Roger Peirce. 25 October 2005 and 14 Dec 2005, telephone.

Paul Pittenger. Worked for CompuServe during the Point of Sale Dial Terminal trials. 17 July 2006, San Mateo, CA.

Bill Powar. 11 November 2005, 16 November 2005, and 23 January 2006, Palo Alto, CA.

Bill Reid. 3 November 2004 and 5 November 2004, Denver, CO.

Brian Ruder. 16 May 2006, telephone.

Chuck Russell. 14 October 2005, telephone.

Ron Schmidt. 30 May 2006, Walnut Creek, CA.

Diderik Schonheyder. 17 October 2006, telephone.

Tom Schramm. 13 January 2006 and 1 February 2006, Crow Canyon, CA.

Jean Stewart. 9 November 2005, San Mateo, CA.

Dawn Tindal. 12 May 2006 and 24 May 2006, telephone.

Aram Tootelian. 6 March 2006, telephone.

John Totten. 2 March 2006 and 3 March 2006, telephone.

B Ray Traweek. 12 May 2006 and 24 May 2006, Roseville, CA.

Pete Yeatrakas. President of WesPay, formerly the California Automated Clearinghouse Association. 19 December 2006, telephone.

List of Acronyms

ABA American Bankers Association, the professional organization that represents the interests of bankers in America.

ACH Automated Clearing House (or Automated Clearinghouse). A computerized system for clearing transactions submitted in electronic form.

ACP Airline Control Program, the IBM operating system and database originally created for the Sabre system, and used for Visa's authorization system since 1977. Now known as Transaction Processing Facility.

ANSI American National Standards Institute.

BASC BankAmericard Service Corporation, a subsidiary of Bank of America created to administer the BankAmericard licensing program in 1966. This organization was replaced for the domestic members in 1970 by NBI, and for the international members in 1974 by IBANCO.

BASE Originally BankAmericard Authorization System Experimental, later changed to BankAmericard Service Exchange, and today treated simply as a title for Visa's electronic authorization and clearing and settlement systems (BASE I and II respectively).

BWG Blue-White-and-Gold, often used to describe the bands design that was part of the Visa mark until recently.

CMA Cash Management Account, a hybrid between an investment and demand deposit account featuring a money market fund attached to an investment portfolio.

CNBT City National Bank and Trust of Columbus, Ohio.

CSI Credit Systems Incorporated, the organization that built and operated the first authorization system for the Interbank system.

DEC Digital Equipment Corporation, suppliers of the PDP line of computers used for NBI's first authorization system.

DOJ The US Department of Justice.

ECR Electronic Cash Register.

EFT Electronic Funds Transfer.

EFTS Electronic Funds Transfer System, a system capable of transferring funds in electronic form.

EIRF Electronic Interchange Reimbursement Fee, a discounted fee available for transactions authorized electronically and posted within three days.

ESBA Eastern States Bankcard Association, the processor for several Interbank members on the East Coast.

FNCB First National City Bank, known today as Citibank.

IATA International Air Transport Association.

IBANCO International Bankcard Company, the original name for the international version of NBI, later changed to Visa International.

INAS Interbank National Authorization System.

ISO International Organization for Standardization. Technically ISO is not an acronym, but a word derived from the Greek *isos*, meaning equal. See `http://www.iso.org/iso/en/networking/pr/isoname/isoname.html`.

MAPS The ABA Monetary and Payments System planning committee.

MICR Magnetic Ink Character Recognition, a technique for encoding the issuing bank, account, and amount on a paper draft, which can be read by both humans and machines.

NBI National BankAmericard Incorporated, the original name of the organization now known as Visa. The name was formally changed to Visa in 1977.

NDC National Data Corporation, a large processor for both Interbank and BankAmericard members in the 1970s.

OCR Optical Character Recognition, a technique for machine-reading printed or hand-written characters.

PARS Programmed Airline Reservation System, the commercial name for the repackaged Sabre system.

PERT The Program Evaluation and Review Technique, a project management technique used to understand dependencies between tasks.

POS Point of Sale. This is commonly used as a modifier for "terminal" when the terminal is designed to be used at the point of sale.

RFP Request for Proposal, a document sent to potential contractors describing the system desired by the client.

SRI Originally Stanford Research Institute, but this organization formally changed their name SRI International in 1977.

TIRF Terminal Interchange Reimbursement Fee, the discounted fee paid by acquirers to issuers for transactions authorized through a point of sale terminal.

TPF Transaction Processing Facility, the current name for Airline Control Program.

TTI Transaction Technology Incorporated, a subsidiary of Citicorp that developed electronic transaction processing technologies.

TTU Tape Transmission Unit, the name given to the DEC computers and magnetic tape readers installed in NBI member processing locations for use with the BASE II system.

UATP Universal Air Travel Plan, a credit card program run by the airlines.

VISA Visa International Services Association. This is the formal name of the Visa organization, though it is commonly referred to simply as "Visa" after first use.

WATS Wide Area Telephone Service, a fee structure that allows organizations to provide toll-free numbers to consumers (IN-WATS), or make long distance calls for discounted rates (OUT-WATS).

WSBA Western States Bankcard Association. The processor for several Interbank members on the West Coast.